Insurance Mathematics and Economics 120 (2025) 107–130

Contents lists available at ScienceDirect

Insurance: Mathematics and Economics

journal homepage: www.elsevier.com/locate/ime

Evolution of institutional long-term care costs based on health factors

Aleksandr Shemendyuk [a,*], Joël Wagner [a,b]

[a] Department of Actuarial Science, Faculty of Business and Economics (HEC Lausanne), University of Lausanne, Extranef, 1015 Lausanne, Switzerland
[b] Swiss Finance Institute, University of Lausanne, Switzerland

ARTICLE INFO

JEL classification:
G22
I13
I18
C33
C55

Keywords:
Long-term care
Institutional care
Costs of care
Multi-state Markov modeling
Empirical data

ABSTRACT

As many developed countries face the challenges of an aging population, the need to efficiently plan and finance long-term care (LTC) becomes increasingly important. Understanding the dynamics of care requirements and their associated costs is essential for sustainable healthcare systems. In this study, we employ a multi-state Markov model to analyze the transitions between care states of elderly individuals within institutional LTC in the canton of Geneva, Switzerland. Utilizing a comprehensive dataset of 21 494 elderly residents, we grouped care levels into four broader categories reflecting the range from quasi-autonomy to severe dependency. Our model considers fixed covariates at admission, such as demographic details, medical diagnoses, and levels of dependence, to forecast transitions and associated costs. The main results illustrate significant variations in care trajectories and LTC costs across different health profiles, notably influenced by gender and initial care state. Females generally require longer periods with less intensive care, while conditions like severe and nervous diseases show quicker progression to more intensive care and higher initial costs. These transitions and expected length of stay in each state directly impact LTC costs, highlighting the necessity of advanced strategies to manage the financial burden. Our findings offer insights that can be utilized to optimize LTC services in response to the specific needs of institutionalized elderly people. These findings can be applied to enhance healthcare planning, the preparedness of infrastructure, and the design of insurance products.

1. Introduction

The demographic shift toward an aging population poses significant challenges to long-term care (LTC) systems worldwide. As life expectancy increases, so does the prevalence of age-related health problems, necessitating expanded services and resources to support the elderly, particularly in their activities of daily living (ADL). Studies like those by OECD (2017) and Kempen et al. (1995) have highlighted the growing demand for healthcare services as more individuals live into their later years, often accompanied by complex health conditions such as multiple diseases (van den Akker et al., 1998) which amplify the need for continuous care (Stark et al., 1995).

In this context, institutional LTC emerges as a critical component of elder care, designed to support those who require substantial assistance. Unlike home or family-based care, institutional settings provide organized, comprehensive care that integrates medical, personal, and social services in a single facility. However, this system also involves significant challenges in terms of financing (Brown and Finkelstein, 2009), availability of care facilities (Katz, 2011), and the recruitment and training of professional caregivers (Nichols et al., 2010). The integration of effective management strategies and sustainable financing solutions is essential to prepare for the coming increase in demand, underscoring the importance of detailed analysis and strategic planning in LTC provision (Colombo et al., 2011; Cosandey, 2016).

Research on LTC costs highlights the significance of modeling in understanding and predicting the financial implications associated with varying durations of care and intensities of service provision. The economic burden on LTC facilities is primarily determined by the length of stay of residents, which varies based on demographic factors, medical conditions, and the severity of physical and psychological impairments (Mathers, 1996; Deeg et al., 2002; Germain et al., 2016). Works by Bladt et al. (2023) and Shemendyuk and Wagner (2024) have shown how age, gender, and specific health profiles influence the demand for care and, consequently, the costs incurred. Particularly, individuals with complex

* Corresponding author.
E-mail addresses: aleksandr.shemendyuk@unil.ch (A. Shemendyuk), joel.wagner@unil.ch (J. Wagner).

https://doi.org/10.1016/j.insmatheco.2024.11.007
Received 17 June 2024; Received in revised form 10 October 2024; Accepted 11 November 2024
Available online 19 November 2024

health conditions such as musculoskeletal and osteoarticular disorders often have extended stays due to lower mortality rates (Makam et al., 2019). Moreover, the intensity of care, measured by the daily time nurses spend with patients, directly impacts the cost structure within LTC settings (Dorr et al., 2005). Studies such as Guccione et al. (1994) and Fong (2019) have shown how the level of dependency due to multiple morbidities increases the need for more intensive and frequent care interventions, thereby escalating the overall costs. This correlation is further complicated by impairments in psychological and sensory functions, which necessitate higher levels of assistance and lead to greater dependency (Marengoni et al., 2011; Barnett et al., 2012).

In LTC cost analysis, multi-state modeling plays a critical role in mapping the complex relationships between health conditions and care trajectories. The development of semi-Markov models, as explored by Fuino and Wagner (2018), enhances understanding of the care paths essential for elderly care management and the design of tailored insurance products. These models effectively track the transitions between different states of health, which are directly influenced by the severity of conditions and determine the duration of stay and intensity of care required (Fong et al., 2015; Sherris and Wei, 2021). The actuarial assessment of LTC products often relies on such models, as they allow for the estimation of transition probabilities that are not only dependent on the current health state but also on the duration within that state, providing a more nuanced view of care dynamics (Pritchard, 2006; Christiansen, 2012; Haberman and Pitacco, 2018). Historically, these models have been used to determine insurance premiums and manage risk by considering both the progression of the health status and its implications on care needs (Govorun et al., 2015; Ai et al., 2017). Studies like those by Czado and Rudolph (2002) and Helms et al. (2005) have extended traditional Markov models to incorporate time-dependent variables, which significantly impact the calculation of costs in LTC settings. This semi-Markov approach, recognized for its ability to integrate time-dependent transitions, offers a sophisticated framework for predicting LTC costs by accounting for the complexity of health trajectories and the direct impact of functional disabilities on life expectancy and subsequent care requirements (Janssen and Manca, 2001; Foucher et al., 2010).

Using a multi-state model, our study aims to analyze the evolution of individual health and its implications on institutional LTC needs and their financing in the context of Switzerland. By leveraging the Swiss social health insurance system's categorization of LTC needs into twelve levels, our model captures transitions between different states of care, including the absorbing state of death. We utilize a comprehensive panel dataset from the LTC institutions of the Canton of Geneva covering the years from 1996 to 2018, which includes detailed records of 21 494 individuals collected using the Canadian “PLAISIR” method (Roussel and Tilquin, 1993). We estimate transition probabilities and associated costs linking them to the individual characteristics known at admission in the institution. This methodology aids nursing staff by predicting care requirements from initial health assessments, supports infrastructure planning by forecasting occupancy, and informs both public and private insurers about expected costs. The latter is essential not only for designing social health insurance policies but also for developing novel private insurance products.

By analyzing various health profiles, our study suggests that the baseline health profile, most commonly observed among institutionalized elderly people, incurs higher LTC costs due to extended care needs stemming from prolonged survival times. Conversely, profiles characterized by severe conditions and nervous diseases demonstrate swift progression to higher dependency states, accumulating significant costs early on, especially among females. Another notable finding is that individuals with cerebrovascular conditions experience a slower progression to severe states yet eventually accumulate substantial costs. Moreover, the tumor disease profile uniquely displays rapid transitions to death, yielding the lowest overall costs due to the shortened duration of care.

The remainder of this paper is structured as follows. Section 2 develops a multi-state model for panel data to assess the changes in the health status of institutionalized elderly and the impact on LTC costs within the Swiss social health insurance framework. Section 3 presents our dataset and statistical analysis, emphasizing the advantages of using medical evaluations over traditional survey-based data. Section 4 applies the developed multi-state model, discussing the transformation of variables, model fitting, and examining transition probabilities and associated costs across various health profiles. Finally, Section 5 provides conclusions, summarizing the insights obtained from our analysis and suggesting directions for future research.

2. Modeling insured LTC costs: framework and methodology

In this section, we develop a model to assess the changes in the health status of institutionalized elderly and the effect on LTC costs within the Swiss social health insurance framework. We start with an overview of the care classification and the reimbursement levels in Switzerland. Next, we introduce the individual's evolution of care and formulate a time-homogeneous multi-state Markov model that describes the underlying process. Then, we detail the likelihood function and the role of initial covariates in determining transition intensities. Finally, we describe the calculation of key metrics, such as transition probabilities and expected length of stay in the care states, that are essential for estimating the costs of care.

Swiss social health insurance reimbursement scheme. The cost of institutional LTC is significant and, in Switzerland, directly related to the daily care needs of the elderly. While housing costs are borne out-of-pocket by the individuals, Swiss social health insurance reimburses care costs along a twelve-level classification, each level of care needs correlating to specific reimbursement amounts as described by the Federal Department of Home Affairs (2016, Section 3, Art. 7 and 7a). This approach ensures that the financial compensation for LTC is systematically organized, making it directly proportional to the intensity of care required.

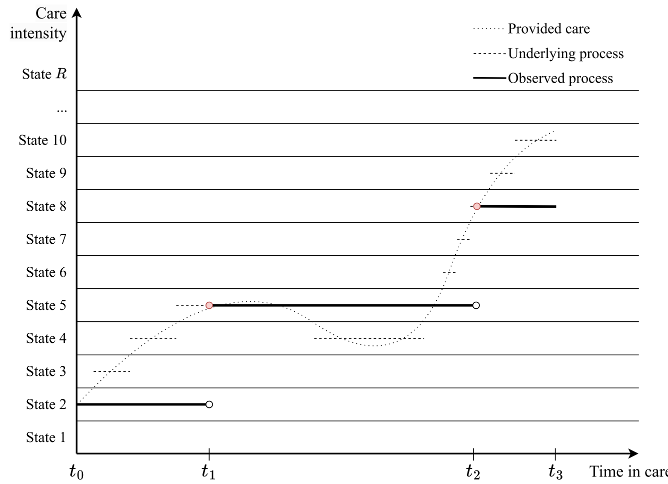
In Switzerland, reimbursement for LTC is determined by twelve ordered levels based on daily care requirements. Under this system, social health insurance pays out daily amounts based on the required level of care:

$$\text{Payout}(r) = 9.60 \times r, \quad \text{(in CHF)}, \tag{1}$$

where $r = 1, 2, \ldots, 12$ denotes one of the twelve categories derived from the minutes of required care per day. These categories start with up to 20 minutes per day, represented by the index $r = 1$ and coming with costs of CHF 9.60. The categories increase by 20 minutes per day for the next states $r = 2, \ldots, 11$. For example, an elderly person requiring 21 to 40 minutes of care per day is represented by category $r = 2$ and the costs yield CHF 19.20. This pattern continues until the final category, $r = 12$, which represents 220 or more minutes of care per day and yields costs of CHF 115.20.

Comparable cash-for-care schemes exist in other European countries, where LTC insurance benefits are structured in several tiers, similar to the Swiss model's categorization of dependency levels. Countries such as Austria, France, and Germany have developed systems that reflect different levels of dependency, similar to the Swiss categorization of care needs. Da Roit and Le Bihan (2010, see Table 1) provide a comprehensive analysis of the European landscape, highlighting the differences in schemes and the funding systems in countries such as Sweden, Netherlands, France, Germany, Austria, and Italy, and their respective financial implications. Furthermore, Yang et al. (2016) examined China's approach to LTC financing, revealing diverse strategies such as Shanghai's social health insurance, Qingdao's LTC insurance, and Nanjing's means testing. Despite differences in healthcare integration and government funding reliance, these models share a core objective with their European counterparts, namely, to provide an adequate reimbursement scheme for the institutional LTC.

Multi-state model framework. In the following, we consider an insurance reimbursement scheme that pays for provided care based on



Note: The dotted curve in graph (a) represents the continuous evolution of care provided at the institution, while the dashed and solid lines represent the underlying and observed processes corresponding to discrete states of care reimbursements.

Fig. 1. Illustration of the care intensity path over time.

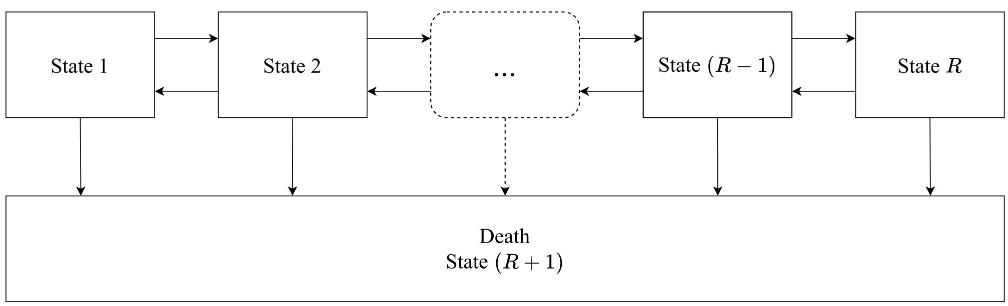


Fig. 2. Transitions of the underlying process in the LTC multi-state model.

R categories. Assuming continuous evolution of the provided care, the payout levels are evolving as a discrete-space continuous-time jump process. In Fig. 1, we illustrate a sample path of care intensity over time, i.e., the evolution of a person's care needs since admission to the institution. The *provided care* is denoted by the dotted curve and represents the continuous evolution of care provided to the elderly. The *underlying* and *observed processes* represented by the dashed and solid lines, respectively, correspond to the R reimbursement levels that follow the multi-state process. The underlying process is directly related to the provided care so that the corresponding multi-state process evolves from one neighboring state to another. Also, it is possible to transition to the absorbing state denoted as "Death" at any point. Fig. 2 shows the diagram with the possible transitions between the model's states. Unlike the underlying process, the observed process represents the administrative data collection procedure that starts at the date of admission $t_0 = 0$ and, in general, is carried out periodically at undetermined times t_1, t_2, and so on. The last observation in time t_3 illustrated in Fig. 1 can indicate the person's moment of death or correspond to the current length of stay in the institution (e.g., related to the end of the observation period due to data extraction). In the latter case, the duration until the next state transition (time-to-event) remains undetermined, and the health state at the date of data extraction is therefore unknown (see the mismatch between the underlying and observed processes in time t_3).

In our analysis, we aim to apply a multi-state Markov model on panel data. For doing so, we consider a framework consisting of $(R+1)$ states, where each state $r = 1, 2, \dots, R$ indicates different care needs, and the state $(R+1)$ denotes the terminal state of death. The transition intensities $q_{rs}(\mathbf{z})$ measure the instantaneous probability of transitioning from state r to state s, for $r, s = 1, \dots, R+1$, $r \neq s$, and are independent of the process history under the Markov assumption (Cox and Miller, 1965). These transition intensities are contained in a matrix Q of dimension $(R+1) \times (R+1)$ with the rows summing up to zero, i.e., the diagonal elements are defined as $q_{rr}(\mathbf{z}) = -\sum_{s \neq r} q_{rs}(\mathbf{z})$. The model allows only for transitions between neighboring states and to the absorbing state so that the matrix Q has the following form:

$$Q = \begin{pmatrix} -q_{12}-q_{1,R+1} & q_{12} & 0 & \cdots & 0 & q_{1,R+1} \\ q_{21} & -q_{21}-q_{23}-q_{2,R+1} & q_{23} & \cdots & 0 & q_{2,R+1} \\ 0 & q_{32} & -q_{32}-q_{34}-q_{3,R+1} & \cdots & 0 & q_{3,R+1} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & -q_{R,R-1}-q_{R,R+1} & q_{R,R+1} \\ 0 & 0 & 0 & \cdots & 0 & 0 \end{pmatrix} \tag{2}$$

In a transition probability matrix $P(t)$, the element $p_{rs}(t)$ represents the probability of an individual transitioning from state r to state s over time t, assuming a time-homogeneous Markov process. The matrix $P(t)$ is defined by the matrix exponential of Q scaled by the time interval t, i.e.,

$$P(t) = \text{Exp}(tQ). \tag{3}$$

This matrix is crucial for our analysis of care trajectories as it helps to assess the expected length of stay in each care state and thus facilitates the cost evaluations.

Likelihood for panel data. To calculate the maximum likelihood estimate of the transition intensity matrix Q, Kalbfleisch and Lawless (1985) and Kay (1986) established a method for a general multi-state model in continuous time with an arbitrary transition matrix $P(t)$. In this context, $i = 1, \ldots, M$ indexes unique trajectories of M individuals through various care states over time. The indices i and j of the function $\mathcal{L}$ represent the likelihood contribution from the j-th transition of the i-th individual in terms of the transition probability matrix. Here, j represents a specific transition event for that individual, moving from one observed state to another over a discrete interval. For intermittently observed processes, the likelihood contribution for individual i from a pair of successive observed states $S(t_j)$ and $S(t_{j+1})$ is given by:

$$\mathcal{L}_{i,j} = p_{S(t_j)S(t_{j+1})}(t_{j+1} - t_j),$$

where $p_{rs}(t)$ denotes the probability of transitioning from state r to state s over time t, derived from the transition probability matrix $P(t)$.

In cases where the times of death are exactly known, the contribution to the likelihood accounts for the uncertainty in the state just before death by summing over all potential states s preceding the terminal state of death:

$$\mathcal{L}_{i,j} = \sum_{s=1}^{R} p_{S(t_j),s}(t_{j+1} - t_j) \cdot q_{s,(R+1)}.$$

In panel data that is limited in time, some individuals are observed to reach the absorbing state, while others are still alive at the end of the observation period, with their most recent health state recorded. The subsequent transition for the surviving individuals, whether to another care state or death, is not observed, leading to right-censoring. For the likelihood calculation, this scenario requires accounting for both the certainty of the last known state and the potential for any future state transitions. In this context, n_i denotes the index of the last observation for individual i. The likelihood for transitions from this last observed state includes all possible subsequent states excluding death and is represented as:

$$\mathcal{L}_{i,n_i} = \sum_{s=1}^{R} p_{S(t_{n_i}),s}(t_{n_i+1} - t_{n_i}).$$

The total likelihood $\mathcal{L}(Q)$ of the multi-state model is constructed by multiplying all individual likelihood contributions $\mathcal{L}_{i,j}$ across every transition and for each individual in the study:

$$\mathcal{L}(Q) = \prod_{i,j} \mathcal{L}_{i,j}. \tag{4}$$

Effect of covariates. In our analysis, we aim to consider the effect of covariates $\mathbf{z}$ that are valued at the time of entry into institutional LTC and do not evolve over time. Incorporating fixed covariates simplifies the estimation of future care costs, even under uncertainty about future health outcomes. This is also consistent with practical needs for predicting care trajectories at the time of entry for new patients. According to Marshall and Jones (1995), the transition intensities q_{rs} can be modeled as functions of these covariates using a proportional hazards method:

$$q_{rs}(\mathbf{z}) = q_{rs}^{(0)} \exp(\boldsymbol{\beta}_{rs}^T \mathbf{z}), \tag{5}$$

where $\mathbf{z}$ represents the vector of covariates fixed at entry, $\boldsymbol{\beta}_{rs}$ is the vector of coefficients associated with the covariates $\mathbf{z}$ for the transition from state r to state s, and $q_{rs}^{(0)}$ is the baseline transition intensity as defined in the matrix Q above. Consequently, incorporating these covariates into the transition intensities influences the total likelihood function. The process of finding optimal values involves maximizing the likelihood function $\mathcal{L}(Q)$ in Equation (4) with respect to $q_{rs}^{(0)}$ and $\boldsymbol{\beta}_{rs}$.

Model output. Once the model parameters are estimated from the data, we compute the key metrics of interest. Specifically, we want to evaluate the probability of transitioning to a state by a given time t, denoted in Equation (3), and the average time an individual starting in state r is expected to spend in each state $s = 1, 2, \ldots, R, R+1$, including death, up to time t:

$$E_{rs}(t, \mathbf{z}) = \int_0^t p_{rs}(u, \mathbf{z})\, \mathrm{d}u. \tag{6}$$

The latter allows for estimating the average care costs over a specified period. Given the payout amounts from Equation (1), the average cost of an institutionalized individual starting in state r with initial covariates $\mathbf{z}$ over time t can be described as:

$$C_r(t, \mathbf{z}) = \sum_{s=1}^{R} E_{rs}(t, \mathbf{z}) \cdot \mathrm{Payout}(s). \tag{7}$$

Here, we consider the average costs as the average number of days spent in a particular state by the time t multiplied by the daily cost of the states. In the $(R+1)$-th state, representing death, no cost arises. However, in the case of modeling a mixed insurance product, a lump-sum term representing a one-time death benefit could be added.

3. Dataset and descriptive statistics

In this section, we present the main characteristics of our dataset and statistical analysis. Section 3.1 provides an overview of our dataset, which offers several advantages over typical survey-based datasets commonly used in LTC research.[1] These advantages are based on the medical evaluations of an individual's health compared to self-reported data, and consistent follow-up during the study. This enables a more precise examination of the health transitions and care requirements within the institutionalized elderly population, overcoming the common limitations of uncertain times of transitions between states and imprecise health reports. Next, in Section 3.2, we analyze the health evaluations recorded in our dataset and the observed transitions among different care states. After consolidating the twelve available care levels into four broader categories, we utilize the Aalen-Johansen estimator to calculate state occupancy over time, enhancing our understanding of care dynamics. Additionally, we stratify these estimates by key covariates such as gender, medical diagnoses, and levels of dependence and provide an analysis of the associated LTC costs.

3.1. Description of the data

This study is based on the private dataset from nursing homes in the Canton of Geneva, Switzerland, provided by the Republic and Canton of Geneva, General Directorate for Health (2019), which was previously studied by Bladt et al. (2023) and later by Shemendyuk and Wagner (2024). The dataset includes $M = 21\,494$ individuals aged 65 or older,[2] consisting of 17 832 complete observations of individuals who died during the study period and 3 662 right-censored observations of those still alive at the time of data extraction.[3] This study covers the period

[1] See for example, the Health and Retirement Study in the United States (HRS), originally reviewed by Juster and Suzman (1995) and later by Sonnega et al. (2014) and Fisher and Ryan (2017), available at https://hrs.isr.umich.edu/; the Survey of Health, Ageing and Retirement in Europe (SHARE), see https://share-eric.eu/ and its introduction by Börsch-Supan et al. (2013); and the China Health, Aging, and Retirement Longitudinal Study (CHARLS) from https://charls.pku.edu.cn/en/ explained by Zhao et al. (2012).

[2] The reduction in the number of observations from Shemendyuk and Wagner (2024) is due to additional quality checks that were implemented when incorporating subsequent health evaluations for each individual. Specifically, we excluded 8 individuals due to discrepancies between the recorded number of health evaluations and the value registered in the personal summary data field. Additionally, 27 and 20 individuals were excluded due to incorrect or inconsistent entry or exit dates, respectively.

[3] Right-censored observations refer to those individuals whose health state is recorded from their entry into the institution until the last observed health

Table 1
Description of the variables.

Variable	Description	Values
Basic information on the pathways		
M	Number of individuals in the dataset	21 494
i	Index of an individual	$1, 2, 3, \ldots, M$
n_i	Number of health evaluations for individual i during the study period	$1, 2, 3, \ldots$
j	Index of the health evaluation for individual i	$0, 1, 2, \ldots, n_i$
t_{ij}	Time of the health evaluation j for individual i	$1, 2, 3, \ldots$ (in days after admission)
$T_{t_{ij}}$	Intensity of care provided per week observed at time t_{ij} for individual i at their health evaluation j	number of minutes (between 0 and 10 080)
$r_{t_{ij}}$	Care level derived from daily care observed at time t_{ij} for individual i at their health evaluation j	$1, 2, 3, \ldots, 12$ (categorical)
Demographic variables		
AG	Age at entry in the institution	$65, 66, 67, \ldots$ (in years)
GE	Gender	female, male (binary)
Medical diagnoses		
ND	Number of diagnoses	$1, 2, 3, \ldots, 9$
$D1$	Diagnosis of first importance	mental, cerebrovascular, respiratory, blood, nervous, osteoarticular, endocrine, heart, tumors, other (categorical)
Di	Diagnosis of k-th importance, $k = 2, 3, \ldots, 9$	see $D1$, plus "none"
Levels of dependence		
DP	Dependence in ADL	$1, 2, 3, \ldots, 9$ (categorical)
PM	Physical mobility limitations	"
OR	Orientation problems	"
OC	Occupational limitations	"
SI	Social integration limitations	"
Impairments of psychological and sensory functions		
VS	Vision	adequate, mild, moderate, severe
HR	Hearing	"

Note: *Only two of the 16 available impairments of psychological and sensory functions appear in our model after the variable selection procedure (see Section 4.1).

from 1996 to 2018 and is collected using the EROS assessment tool, a methodology developed by Roussel and Tilquin (1993). In our dataset, all institutionalized individuals have no instances of leaving and reentering the institution, thus, providing consistent tracking of their LTC pathways.

Our data captures several categories of variables: demographic information, medical diagnoses, levels of dependence, impairments of psychological and sensory functions, and the intensity of care. The latter quantifies the care provided to an individual over a one-week health evaluation period and is measured in minutes of care per week. Based on this variable and within the Swiss reimbursement scheme, we derive one of the twelve cost levels according to the daily care requirement, see Federal Department of Home Affairs (2016) and Section 2. Table 1 highlights the variables related to the pathway of elderly people receiving institutional LTC. Further, we borrow parts of the explanation of the other available variables from Bladt et al. (2023, Section 3.1) and Shemendyuk and Wagner (2024, Section 2.1) and provide specific details where needed. Since we account for multiple health evaluations per individual, we provide more details on the intensity of care variable and its related reimbursement level. Furthermore, we introduce the observed time spent in a state.

Pathway variables. For each individual $i = 1, 2, \ldots, M$, our data records their stay in institutional LTC from admission until death, if applicable, or the date of the data extraction, August 21st, 2018. Upon entry into the facility, every individual undergoes a detailed initial health screening lasting for one week, initializing the start of their care path. This initial screening, indexed as $j = 0$ and time $t = 0$, forms a baseline of health information, including medical diagnoses, levels of dependence, and impairments of psychological and sensory functions, alongside the intensity of care.

The subsequent health evaluations $j = 1, 2, \ldots, n_i$ of an individual i are periodically conducted at random intervals, typically ranging from one to two years. These evaluations update each individual's health information, reflecting changes in their care needs. Each health evaluation is indexed by j, denoting the evaluation sequence for an individual, and the specific time t_{ij} when the evaluation was conducted, recorded in days from the initial entry into the institution. The intensity of care $T_{t_{ij}}$, observed during these evaluations, is measured in minutes of care provided per week at each time point t_{ij}. Dividing $T_{t_{ij}}$ by seven gives the daily intensity of care and indicates the level of care $r_{t_{ij}}$, corresponding to one of the 12 ordered levels of the Swiss reimbursement scheme. These categories reflect the range of care needs from minimal assistance to extensive care requirements, with higher numbers indicating a greater need for daily care. By the end of the study, each individual's care path is characterized by the number of health evaluations n_i, including the initial screening at entry.

We update the intensity of care and the corresponding level of care throughout subsequent health evaluations while keeping the values from the initial assessment for the other covariates (see below). This approach limits the model's complexity and enables a straightforward prediction of LTC costs based on the initial values of the covariates.

Demographic variables. The demographic characteristics of the individuals in our study are primarily defined by the age at entry into the institution (AG) and the gender (GE). The age at entry is computed based on the date of birth and the date of admission, reflecting the full years that have passed until the entry into institutional LTC. Our dataset

evaluation but where the recording is interrupted by the end of the observation period. Thus, the duration until the next transition (time-to-event) remains undetermined, and similarly, the health state at the time of data extraction is not recorded.

consists of a broad age range at entry, from 65 years, ensuring that all individuals are of retirement age or older, to the oldest recorded entry at 106 years. Gender is identified as a binary factor, distinguishing between "male" and "female" categories.

Medical diagnoses. Our dataset includes up to nine medical diagnoses ($D1, D2, \ldots, D9$) for each individual, with $D1$ representing the primary condition and the others ranked by decreasing importance. If an individual has fewer than nine diagnoses, subsequent values are assigned as "none." Diagnoses are encoded following the International Classification of Diseases (ICD) standards detailed by the World and Organization (2016), and aggregated further into general groups: mental, cerebrovascular, respiratory, blood, nervous, osteoarticular, endocrine, heart, tumors, and an "other" category for remaining conditions.[4]

Levels of dependence. Dependence levels are evaluated based on five dimensions to measure individuals' varying degrees of physical and social limitations. These dimensions include limitations in ADL (denoted as DP), physical mobility (PM), orientation (OR), occupational activities (OC), and social integration (SI). Following the guidelines established by the World and Organization (1980), these variables are recorded on a nine-point scale as ordered factors that categorize the severity of limitations from minimal to severe with levels 1 to 9, respectively (also see Bladt et al., 2023, Sect. 3.1 and Table 2). Specifically, DP evaluates the individuals' independence in performing both basic ADL, such as personal hygiene, eating, and dressing, and instrumental ADL, like housekeeping and cooking. PM assesses the ability to move effectively within the environment, considering the use of mechanical aids but excluding assistance from others. OR measures cognitive functions related to understanding and interacting with the environment. OC assesses the capacity to engage in customary activities reflecting the individual's age and gender within the institutional setting. Lastly, SI looks at the individuals' ability to participate in social activities and maintain social relationships, which are essential for life in an institutional context.

Impairments of psychological and sensory functions. Health records from this group are detailed across 16 variables, each measured on an ordered four-point scale ranging from adequate to severe. These scales assess the severity of psychological and sensory function impairments, incorporating any compensatory mechanisms the individual may use, such as glasses or medication for psychological impairments, and comparing performance against the normative standards of a healthy person of the same age and gender. The impairments evaluated include recent memory (RM), long-term memory (LM), thinking (TH), perception and attention (PA), consciousness and wakefulness (CW), orientation related to time, person, and space (TP), decision-making (DM), impulses (IM), will and motivation (WM), emotions including feelings and mood (EM), behavior (BH), language (LG), sight (VS), hearing (HR), making oneself understood (SU), and understanding others (OU). A comprehensive overview of the original definitions in Roussel and Tilquin (1993), the descriptions of the levels associated with these variables and their impact on an individual's health profile, is available in (Bladt et al., 2023, Sect. 3.1).

Heath profiles. This dataset was explored in Shemendyuk and Wagner (2024), revealing that institutionalized elderly can be categorized into eight distinct health profiles. In Sections 4.3 and 4.4, we utilize them to examine the impact of covariates on LTC costs. The following qualitative summary presents the dominant characteristics of each profile ordered from the largest to the smallest group:

1. *Baseline health profile*: This is the largest group, mainly comprising older women, characterized by minimal care needs and the longest median length of stay, suggesting relatively better health compared to other groups.
2. *General severe conditions*: Includes individuals with significant mental health challenges and high levels of dependence, requiring considerably more care and exhibiting shorter stays than the baseline profile.
3. *Moderate-severe conditions with nervous diseases*: Features the youngest average age at entry and is distinguished by predominant nervous system pathologies, requiring care levels similar to the previous profile.
4. *Moderate conditions with endocrine diseases*: Unique for its high prevalence of endocrine disorders, this group displays moderate levels of dependence and healthcare needs, positioned between the baseline and more severe profiles.
5. *Moderate conditions with cerebrovascular diseases*: Characterized by notable cerebrovascular issues, this profile exhibits slightly higher dependence and healthcare needs than the endocrine profile.
6. *Moderate conditions with respiratory diseases*: Marked by significant respiratory issues, individuals in this group have moderate care needs and one of the shorter median stays.
7. *Moderate conditions with blood diseases*: This profile includes a notable presence of blood disorders associated with moderate care needs and a relatively short median length of stay.
8. *Moderate conditions with tumor diseases*: This is the smallest group characterized by a high prevalence of tumor-related diseases and the shortest median length of stay.

3.2. Descriptive statistics

In the following, we present descriptive statistics that detail the health pathways of individuals receiving institutional LTC. We identify the transitions between care states and note significant observations, such as the absence of individuals in the lowest care state and the prevalence of high levels of care before death. To manage the model's complexity and allow for robust estimates, we consolidate the various care states into broader categories. These categories range from care levels that indicate autonomy to those that indicate severe dependency. This classification allows us to apply the Aalen-Johansen estimator to evaluate occupancy probabilities and associated costs over time. We stratify further by gender, medical diagnoses, and levels of dependence, and explore the implications of these factors on care progression and costs.

While the dataset contains the observations of 21 494 individual care paths (3 662, 17.0%, of which are right-censored), it counts 54 386 health evaluations, including the initial health evaluation at entry. The complete observations contribute to 45 180 evaluations (83.1%), whereas the right-censored paths contribute to 9 206 evaluations (16.9%). Using successive health assessments, we establish transitions considering two consecutive known states $S(t_j)$ and $S(t_{j+1})$ for all individuals and their paths. For individuals still alive at the end of the study, the last observed state does not lead to another within the study period, so the last transition is marked as "RC," indicating right-censored time-to-event. Table 2 reports the number of observed transitions between the care levels.

From Table 2, we observe that none of the individuals were in the lowest care state ($r = 1$), receiving less than 20 minutes of care per day. This suggests that individuals requiring minimal care either do not enter institutional LTC or their needs are evaluated beyond the lowest care level. Furthermore, state $r = 12$ is the most prevalent final state before death, indicating significant care needs for individuals in the final stages of life.

Furthermore, the statistics in Table 2 reveal a trend in which individuals primarily transition to higher levels of care. For instance, from state $r = 3$, part of the individuals remain on the same care level (891 transitions), and a significant number progress to care states $r = 4$ (722 transitions) and $r = 5$ (363 transitions). When focusing on state $r = 7$, a significant number of people transition to states $r = 8$ (448 transitions)

[4] For details on the definition of the disease groups and adaptations from ICD-9 to ICD-10, see Bladt et al. (2023, Footnotes 7 and 8) and Shemendyuk and Wagner (2024, Footnote 2).

Table 2
Number of observed transitions between care levels and right-censoring counts.

From ...	... to 1	2	3	4	5	6	7	8	9	10	11	12	Death	RC
1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	0	175	226	104	35	19	23	19	7	13	5	5	69	6
3	0	40	891	722	363	193	163	139	119	108	69	99	355	169
4	0	10	244	838	682	436	284	282	252	233	142	185	601	250
5	0	0	43	258	486	444	345	337	349	274	183	249	632	273
6	0	0	14	76	203	357	390	352	379	291	202	284	648	270
7	0	0	7	33	82	160	333	448	465	363	256	354	755	307
8	0	0	6	29	44	53	159	429	731	500	368	574	1 138	315
9	0	1	5	29	29	48	72	215	732	823	613	1 096	1 816	429
10	0	0	3	13	22	30	35	88	321	937	817	1 389	2 360	451
11	0	0	4	6	14	19	17	46	129	395	838	1 550	2 389	389
12	0	0	1	5	11	12	19	27	75	174	501	3 696	7 069	803

Note: The categories from 1 to 12 correspond to the care levels. The abbreviation "RC" stands for right-censored observations corresponding to individuals whose last observed state does not lead to another transition.

and $r = 9$ (465 transitions), while only 282 individuals move to lower levels of care ($r = 3, 4, 5$ or 6). This observation is consistent with findings from the extant literature, such as those by Liddle (1992), which suggest that the health conditions of individuals in LTC settings tend to deteriorate due to factors like inadequate resources and underestimation of disabilities. Therefore, it is generally observed that the care needs of elderly individuals in institutional LTC increase over time, with few improvements resulting in mostly only a one-level decrease in care needs. However, worsening conditions can lead to a significant increase in care needs, up to two or three levels higher from one evaluation to another.

Absolute counts of right-censoring become more important in higher care levels, notably for transitions from states $r = 9, 10, 11$ and 12, reaching 803 right-censored transitions for $r = 12$. This observation underscores the critical need for careful planning of LTC services, as a significant number of individuals continue to require intensive care (see, e.g., Burt et al., 2014). This also underlines the relevance of accounting for right-censoring in our model.

To simplify the analysis, avoid computational challenges, and obtain robust results when constructing a multi-state model with many states and limited data, we group the care levels. This is particularly important in scenarios where the dataset may not support a highly detailed model without risking overfitting, especially when assessing the impact of covariates. Accordingly, we aggregate the care categories into four broader groups as depicted in Fig. 3: state A includes levels $r = 1, 2, 3$, state B encompasses $r = 4, 5, 6$, state C comprises $r = 7, 8, 9$, and state D spans levels $r = 10, 11, 12$. This approach aligns and is comparable with the categorization used in prior studies that assess dependency based on limitations in activities of daily living; see, e.g., Rickayzen and Walsh (2002); Biessy (2015); Fuino and Wagner (2018); Esquível et al. (2021). Here, state A is indicative of quasi-autonomy with less than one hour of care per day, B reflects mild dependency or 1-2 hours per day, and C and D mirror moderate and severe dependency levels, corresponding to 2-3 and 3+ hours per day, respectively. The number of observed transitions for the aggregated groups and their respective proportions are presented in Table 3. Using the care costs defined in Equation (1), we consider the following average care costs in the four groups: CHF 19.20 for state A, CHF 48 for state B, CHF 76.80 for state C, and CHF 105.60 for state D.

To analyze transitions between aggregated care levels, we use the Aalen-Johansen estimator from the `survival` package in R, see Therneau (2024). It allows us to assess the probability of occupying each care state over time and calculate the corresponding care costs, also accounting for covariates. For an initial overview of the dataset, Fig. 4a presents Aalen-Johansen estimates across the four aggregated states. The occupancy probabilities demonstrate a tendency for individuals to transition from lower states to more intensive care levels over time. The initial state distribution indicates that approximately 11.5% of the individuals entered institutional LTC in state A, while 32.9% began in state D. The rise in occupancy for state D at times around 26-28 months is probably due to the combined effect of people starting in lower states and developing higher dependency levels over time, and of those starting in state D tending to have a higher death rate, indicating a pivotal moment for care provision in institutional LTC. Further, we present Aalen-Johansen estimates stratified by gender, first medical diagnosis, and levels of dependence. Table 5 in the Appendix presents the details in numbers as well as the results stratified by age at entry.

Gender. As shown by the Aalen-Johansen estimates stratified by gender in Fig. 4b, the probability of males in all care states generally decreases over time. For females, while the overall declining trend in state occupancy is similar, we observe a more pronounced bump in the probability of being in state D at durations of 26 to 28 months since admission. Upon admission, 39.3% of men are in the highest care state D compared to 26.2% in state C, denoting a 13.1% difference between these two states. In contrast, the distribution among females shows a more balanced initial allocation, with states B, C, and D each accounting for around 28-30%, indicating a relatively uniform spread in care requirements at the time of admission.

Fig. 5 presents the cumulative LTC costs for institutionalized elderly by gender. These costs are derived from the Aalen-Johansen estimates of state occupancy, applied in conjunction with Equation (7) to evaluate the average cost. Here, for a female starting in states A, B, C, and D, with initial probabilities of 12.2%, 28.5%, 28.8%, and 30.5%, respectively (see also Fig. 4b), the mean duration in each state is inferred from the Aalen-Johansen estimates. These durations are then multiplied by each state's average daily costs. The cumulative costs for males are calculated in the same way, taking into account their initial state probabilities of 9.5%, 25.0%, 26.2%, and 39.3% in the states from A to D, respectively. After one year, the cumulative costs are comparable for both genders, with CHF 22 841 for women and CHF 21 960 for men. However, as time progresses, we observe a steeper increase in costs for females than males; by the fifth year, a woman reaches cumulative costs of CHF 86 367 on average compared to CHF 67 635 for a man, and by the tenth year, the costs for females average at CHF 115 350 while males cost CHF 77 543. The majority of these expenses are accumulated from state D, which is the most resource-intensive state. This finding is consistent with the trends observed in the Aalen-Johansen estimates from Fig. 4b, which indicated a quicker progression to higher dependency states and higher mortality among males. Indeed, the higher mortality in men significantly limits the costs when compared to women.

Medical diagnoses. Fig. 11 in the Appendix presents the Aalen-Johansen estimates and the corresponding cumulative costs for individuals with a particular primary diagnosis $D1$ at admission. Individuals with cerebrovascular and nervous conditions predominantly begin in

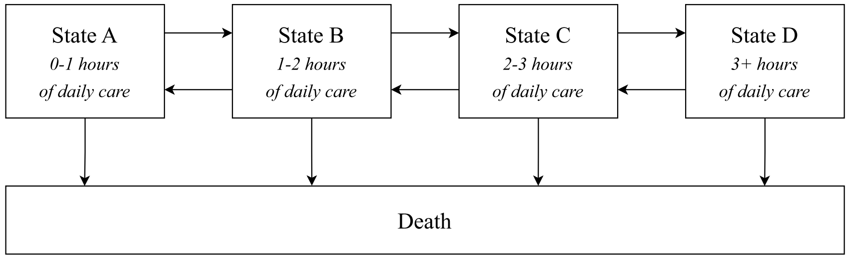
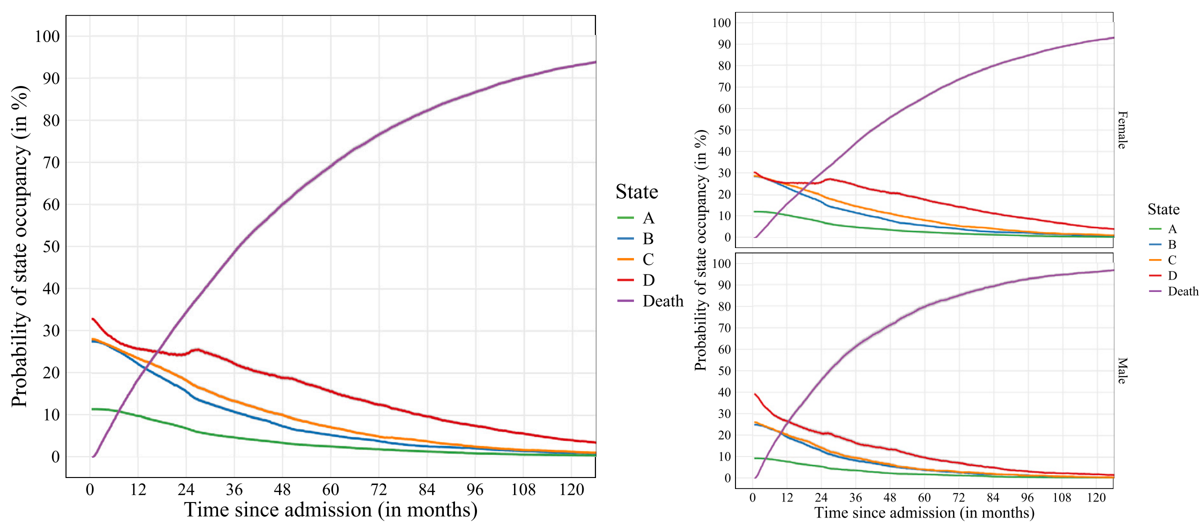


Fig. 3. Transitions of the underlying process in the model with aggregated care levels.

Table 3
Number of transitions between the aggregated care levels and right-censoring counts.

From …	… to A	%	B	%	C	%	D	%	Death	%	RC	%
A	1 332	32.2	1 436	34.7	470	11.4	299	7.2	424	10.3	175	4.2
B	311	2.6	3 780	32.1	2 970	25.2	2 043	17.3	1 881	16.0	793	6.7
C	19	0.1	507	3.7	3 584	25.9	4 947	35.8	3 709	26.8	1 051	7.6
D	8	0.0	132	0.5	757	3.1	10 297	41.8	11 818	47.9	1 643	6.7

Note: States A, B, C, and D represent less than 1, 1-2, 2-3 and 3+ hours of daily care, respectively. The abbreviation "RC" stands for right-censoring. The shares sum up to 100% in each row.



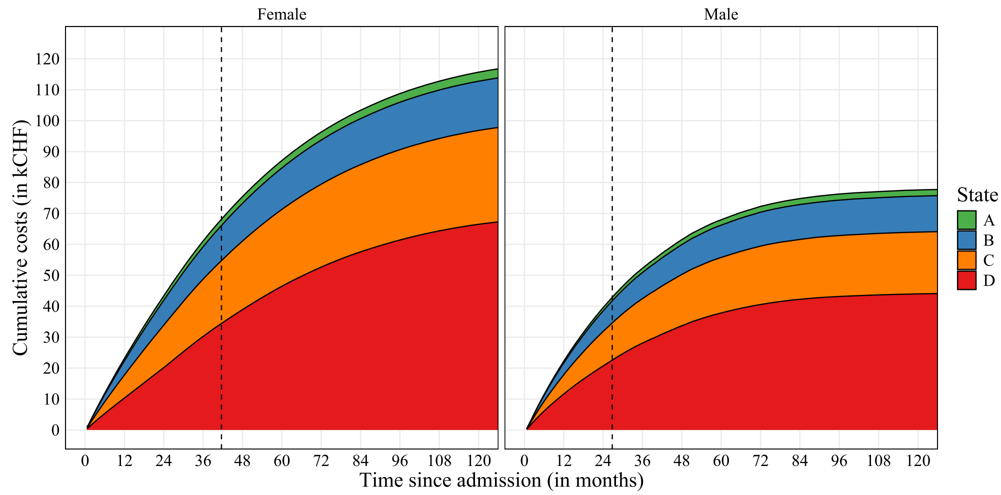
(a) Full dataset. (b) Results by gender.

Fig. 4. Aalen-Johansen estimates with 95% confidence intervals of state occupancy probabilities. (For interpretation of the colors in the figure(s), the reader is referred to the web version of this article.)

higher care states, with approximately 45% and 50% being allocated at admission in states C and D, respectively, reflecting the substantial care needs associated with these diagnoses. In contrast, patients with a mental diagnosis exhibit a more uniform distribution, with around 30% entering states B, C, and D, respectively, which indicates varied care needs at the admission. Patients with osteoarticular, heart, and other conditions show a tendency to start predominantly in state B, suggesting that these conditions are initially present with a relatively mild level of dependency. Notably, those with osteoarticular conditions display lower mortality rates, and the probability of being in state D remains relatively constant at about 20% for up to 56 months after admission. In contrast, individuals diagnosed with tumors have the highest mortality rate, with the median survival time being approximately 8 months.

In terms of cumulative costs, the highest average costs stem from patients with mental and nervous diagnoses, which correlates with their higher needs for care and longer occupation times in state D. Costs for heart disease are more evenly spread across states B, C, and D, suggesting a more balanced progression through the care levels. A similar pattern is observed in patients with osteoarticular diagnoses, who tend to reside in less demanding care states despite longer average lifespans, resulting in lower cumulative costs.

Levels of dependence. Figs. 12–16 in the Appendix present the stratified Aalen-Johansen estimates and cumulative costs across different levels of dependence: dependence from others (DP), physical mobility (PM), orientation (OR), occupation (OC), and social integration (SI). Lower levels in these dependence measures upon admission are related to lower levels of initial care. However, as time progresses, a shift occurs with individuals increasingly transitioning to higher care states C and D. Conversely, those entering LTC with high levels of dependence in any of the five variables predominantly occupy state D, displaying a



Note: The vertical dashed line indicates the median survival time.

Fig. 5. Cumulative 10-year LTC costs by gender based on Aalen-Johansen estimates.

generally consistent decline in survival curves, with notable exceptions. For instance, individuals with a physical mobility (PM) score of 7 and 8 exhibit a significant increase in the probability of being in state D at approximately 26 months after admission. This pattern is also observable in the levels 6 and 7 of the orientation (OR) and social integration (SI) variables. Financially, significant contributions to cumulative costs from state A are primarily seen in those with lower initial levels of dependency. In contrast, for individuals with higher dependency levels, the costs are mainly concentrated in state D, with a lower but still notable portion stemming from state C.

4. Model application and results

When applying the multi-state model described in Section 2, we use the aggregated care levels denoted as states A, B, C, D, and Death as introduced in Section 3.2. For model fitting, we use the `msm` package in R. It is specifically designed for handling panel data (Jackson, 2011). This package supports both numerical and categorical covariates; however, the inclusion of categorical variables significantly increases computational demands due to a sharp rise in the number of parameters that need optimization as seen in Equation (5). Our dataset encompasses a large number of individuals, each with a comprehensive set of health evaluations and numerous variables previously identified as significant in determining care needs and duration of stay in institutional LTC (Bladt et al., 2023; Shemendyuk and Wagner, 2024).[5] To simplify the model fitting, we transform categorical covariates into numerical formats where feasible. We then refine the model by selecting the most relevant variables in our multi-state context. Following these adjustments, the model is analyzed to examine the transition probability matrices for both genders across different ages at entry. This allows us to estimate the average length of stay in each state and calculate the associated costs.

[5] Demographic factors such as age and gender are known to affect the length of stay (Mathers, 1996; Deeg et al., 2002; Germain et al., 2016; Fong et al., 2017; Fuino and Wagner, 2020), while the pathologies, including conditions like musculoskeletal and osteoarticular disorders, influence both stay duration and care intensity (Davidson et al., 1988; Pack, 2009; Makam et al., 2019). Additionally, levels of dependence and impairments in psychological and sensory functions are critical in determining care needs, with multimorbidity leading to increased care burdens (Guccione et al., 1994; Arrighi et al., 2010; Marengoni et al., 2011; Barnett et al., 2012; Koroukian et al., 2016; Albarrán et al., 2019; Fong, 2019; Jennings et al., 2020).

4.1. Data transformation and variable selection

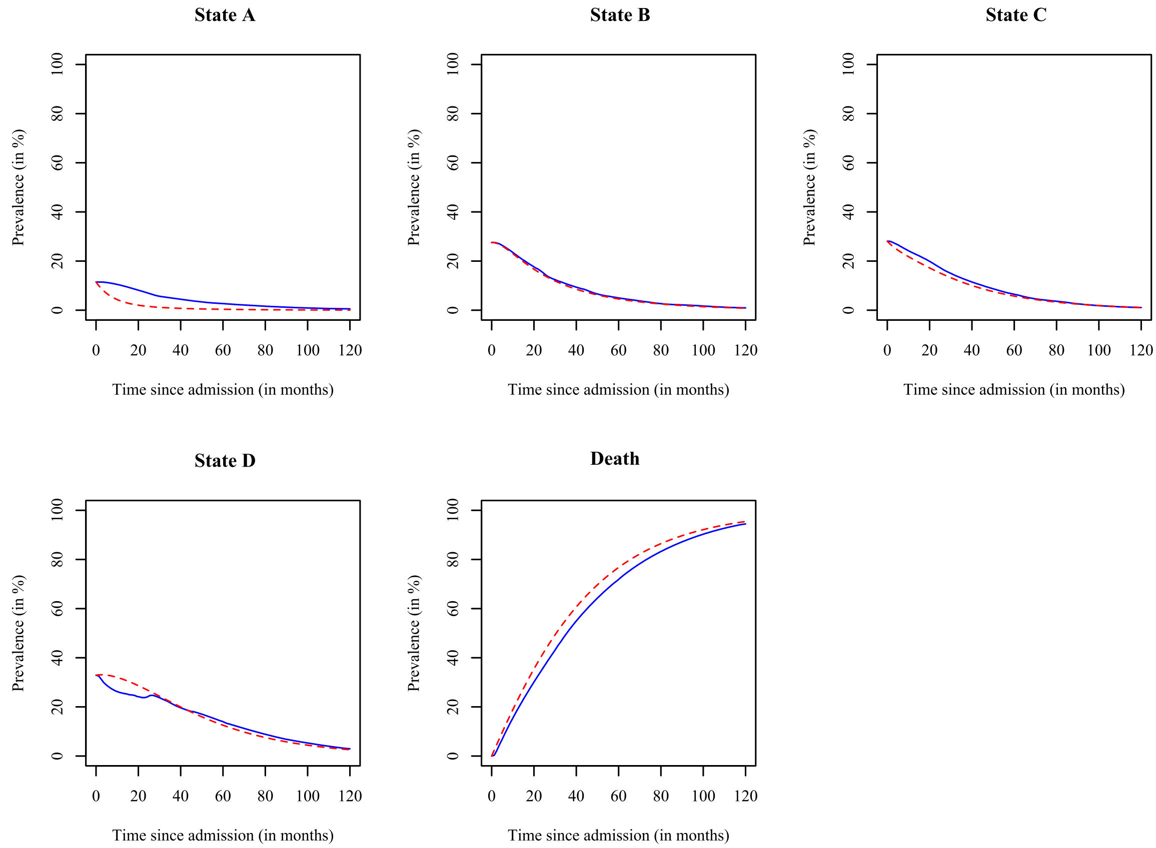
Data transformation. In this paper, we use the methodology detailed in the study by Shemendyuk and Wagner (2024) for calculating pathology scores based on medical diagnoses. It was shown that these scores are critical for assessing the health profiles of the elderly within institutional LTC, which are indicative of the amount of care received and the length of stay. Similarly, the model uses age at entry, levels of dependence, and impairments of psychological and sensory functions as numerics. The exception is the gender variable, which remains binary.

In order to apply the multi-state model to the available data, it is necessary to transform the categorical medical diagnoses, ordered according to importance, into a numerical format. In order to account for the full pathology profile and to maintain the importance ranking of each diagnosis, a score must be calculated for the set of aggregated disease groups: mental, cerebrovascular, nervous, osteoarticular, heart, tumors, and other. The scoring system is adapted as follows:

$$\text{Score}(d) = \sum_{i=1}^{9} (10 - i) \cdot \mathbb{I}(Di = d).$$

Here, d is one of the disease groups, Di is the medical diagnosis at i-th importance rank, and $\mathbb{I}(\cdot)$ is the indicator function. This score is weighted by the rank of the diagnosis, with the first-ranked diagnosis contributing most significantly to the score, and the contribution decreases as the rank lowers.

Variable selection. In order to identify the most influential covariates for our multi-state model, we begin with a null model constructed from the full panel data, which consists of observed transitions among the states without including any covariates. We then apply a stepwise forward procedure based on the Akaike Information Criterion (AIC, see Akaike 1974). In this iterative procedure, one covariate is introduced at a time to the existing model, with the AIC score calculated for each addition. The covariate that yields the greatest reduction in the AIC score is integrated into the model and built upon in the next iteration. This procedure is repeated until the inclusion of new covariates does not result in an improvement in the AIC score. The final model incorporates selected covariates, including age at entry AG, gender GE, number of diagnoses ND, pathology scores for cerebrovascular, nervous, osteoarticular, heart, and tumor diseases, as well as dependency in ADL DP, physical mobility PM, orientation OR, and visual VS and hearing impairments HR.



Note: The solid and dashed lines correspond to the observed prevalence from the data and the expected prevalence from the model, respectively.

Fig. 6. Goodness of fit of the multi-state model for the reference profile.

4.2. Goodness of fit

In this section, we analyze the quality of the multi-state model fit, determining whether the model under- or overestimates the transition probabilities. This ensures that our interpretations of the results are accurate. We introduce a *reference profile* representing the person with the most common values of the covariates in the dataset. That is, regarding the variables selected for the modeling, the reference profile is characterized by an 87-year-old woman with nine medical diagnoses, the most important of which is in the mental category ($D1 = \text{mental}$), followed by eight pathologies $D2, \ldots, D9$ from the "other" group. She is in quasi-permanent need of assistance ($DP = 7$), with mobility limited to the institution ($PM = 6$), and has moderate disorientation ($OR = 5$). Her visual and hearing impairments are classified as "mild". As a result, her score of mental diagnoses is 9, the "other" group leads to a score of 36, while the scores of the remaining pathology groups are zero.

A first indication of the goodness of fit of a multi-state model can be obtained by estimating the observed numbers of individuals occupying a state over a series of times and plotting these against forecasts from the fitted model for each state. Fig. 6 shows the observed share of individuals and the forecasted prevalence rates across all states for an individual corresponding to the reference health profile. The initial probability of being at each state is determined from the data (see Section 3.2).

Across all states, the prevalence estimated with the model generally follows the trends of the observed data, indicating a reasonably good model fit. The discrepancies between observed and expected prevalences appear minimal in states B and C for all times, suggesting a good performance in predicting medium-level care states. However, there are deviations in states A, D, and Death, which could indicate that the effects of certain variables are not fully captured by the model for these states. In particular, the model tends to strongly underestimate the prevalence of individuals in state A in the first five years after admission. This suggests that the model prematurely transitions individuals with the reference health profile to higher dependency states, whereas observations in the data tell that they continue to receive minimal care for much longer periods. This persistent underestimation implies an external factor not captured by the model, influencing the low demand for care. Conversely, in state D, the fit improves significantly after approximately 26 months. However, in the first 26 months after admission, the model overestimates the number of people in this highly care-intensive state. This overestimation seems correlated with a slight underestimation in state C during the same period, suggesting a misclassification of individuals into a higher care state. Additionally, the underestimation in state A could contribute to this early discrepancy in state D. Finally, the model slightly overestimates the probability of death, hinting at additional factors prolonging survival not accounted for in the current model.

It is important to consider that the aggregation of the twelve care levels into four broader categories may contribute to the observed discrepancies between the model's predictions and the observed data. Our decision to group the care levels into states A, B, C, and D was based on practical and clinical considerations, aggregating groups to intuitive and meaningful ranges of care hours per day. This categorization simplifies the analysis and enhances the interpretability of the results for practitioners and policymakers. However, by grouping multiple care levels into broader states, we might be oversimplifying the underlying transition dynamics, potentially masking subtle differences in transition probabilities between individual care levels. This simplification could contribute to the underestimation of prevalence in state A and overestimation in state D, as the model may not fully capture the nuanced progression through the finer-grained care levels. To mitigate this potential misspecification, one could consider increasing the number of states to reflect the original twelve care levels. However, this approach presents significant challenges. Despite our dataset comprising 21 494 individual paths, expanding the number of states would substantially increase the number of parameters to estimate. This approach would require a much larger dataset to avoid overfitting and ensure statistical reliability. Alternatively, advanced modeling techniques could be employed to better capture variability within the aggregated states without increasing the number of states. Methods such as hidden Markov models (Rabiner, 1989), semi-Markov models with time-varying covariates (Titman, 2011), and mixed-effects models (Pinheiro and Bates, 2000) can provide more flexible frameworks to reflect the underlying transition dynamics more accurately.

4.3. Results for the baseline health profile

To better understand the impact of covariates on care trajectories within institutional LTC settings, we analyze the transition probability matrices derived from Equation (3). These matrices reveal how the probabilities of transitioning from one care state to another evolve over time. For categorical covariates, such as gender GE, it is straightforward to visualize differences by comparing side-by-side plots of transition probabilities for females and males. However, this direct comparison approach becomes more complex with numerical covariates, as it would require generating and comparing numerous plots across a spectrum of values for each variable. To avoid this complication and still capture the effects of covariates on care trajectories, we consider the health profiles of elderly individuals in institutional LTC described in Section 3.1. These profiles allow us to illustrate and analyze the expected progression through care states over time, offering insights into how specific covariates influence these transitions.

Baseline health profile. Individuals in this group represent the most common health profile among institutionalized elderly, typically requiring minimal daily care and exhibiting the longest survival times (see the introduction of the health profiles in Section 3.1 and Shemendyuk and Wagner, 2024, Table 3). This group predominantly consists of females (77.5%). It is characterized by a median age at entry of 87 years, with six medical diagnoses in the median, and resulting median diagnosis scores of 5, 7, and 14 for osteoarticular, heart, and other groups, respectively, while scores for other diseases are zero. The median levels of dependence are $DP = 7$, $PM = 6$, and $OR = 5$, with median visual and hearing impairments classified as “mild”.

Fig. 7a illustrates the fitted transition probabilities $P_{rs}(t, \mathbf{z})$ over time t for females with covariates $\mathbf{z}$ of the baseline health profile across the entry ages of 70, 80, and 90 years, stratified by the starting state r (see the label on the right axis indicating the starting state A, B, C, or D in each row of graphs). Females entering the institution at age 70 are more likely to remain in a state with lower care longer than their older counterparts. Notably, at age 90, females demonstrate a higher probability of transitioning directly from state A to state D, indicating a potentially rapid escalation in care needs. These patterns align with Freedman et al. (2016), who found that women accumulate more years in disabled states compared to men, further validating our results.

In contrast, Fig. 7b illustrates the transition probabilities for males. As the age at entry increases, males show a noticeable decline in the likelihood of either remaining in state A or transitioning to it. Conversely, their probability of moving to state D or experiencing death increases, particularly for those entering at higher ages. This finding aligns with research by Sherris and Wei (2021), who used multi-state models to reveal that men typically spend less time in dependent states than women, thus accruing lower long-term care costs overall.

The gender differential in disability transitions is particularly noteworthy. Studies have shown that older women, while living longer than men, spend a greater proportion of these additional years in disabled states. For instance, Chan et al. (2011) demonstrated that females are more likely to remain in disabled states, supporting our findings from this model where females display longer durations in low-care states before transitioning to higher dependency levels. Similarly, Kingston et al. (2014) explored the male-female disability-survival paradox, where women, despite longer life expectancy, tend to experience higher rates of disability transitions due to diseases like cerebrovascular and respiratory conditions.

Fig. 8a displays the average cumulative costs $C_r(t, \mathbf{z})$ by time t for females with covariates $\mathbf{z}$, stratified by age at entry for the four starting states r. In this case, the baseline health profile defines the covariates. Across all age groups, women demonstrate higher cumulative costs than their male counterparts (cf. Fig. 8b). The vertical dashed lines on the graphs represent the median survival times, denoted by δ, that depend on the individual’s starting state r and covariates $\mathbf{z}$. This time marks the duration until the death probability for an individual reaches 50%. In these graphs, the lines are consistently positioned further to the right for females than for males, suggesting that women remain in the LTC system longer. This extended duration contributes to the overall higher costs, as women are more likely to be alive and hence accumulate higher expenses over time. The length of stay is consistently longer for younger individuals and for those admitted to an institution in a state with lower care needs. The prolonged period in state D for females significantly increases the cumulative costs, highlighting the impact of longevity on LTC costs.

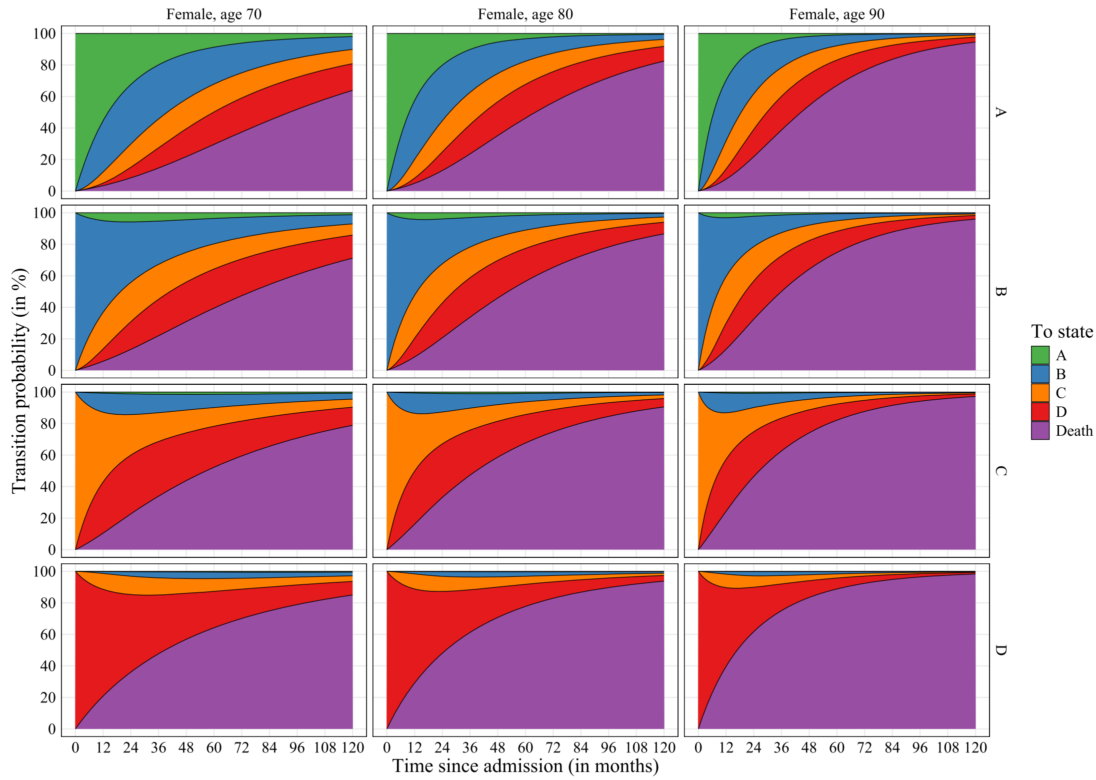
Similarly, Fig. 8b displays the costs for males. The costs exhibit the greatest increase in the initial period following admission, with a pronounced increase for those entering at age 70, indicating that younger males accumulate higher costs at a more rapid pace. The median survival times show that younger entrants reach the transition to death later, corresponding to their higher cumulative costs. Over time, costs associated with state D become more significant for all age groups, emphasizing the financial impact of higher dependency care. However, the costs for state A remain negligible, reflecting its minimal contribution to overall LTC costs for this profile.

4.4. Results for other health profiles

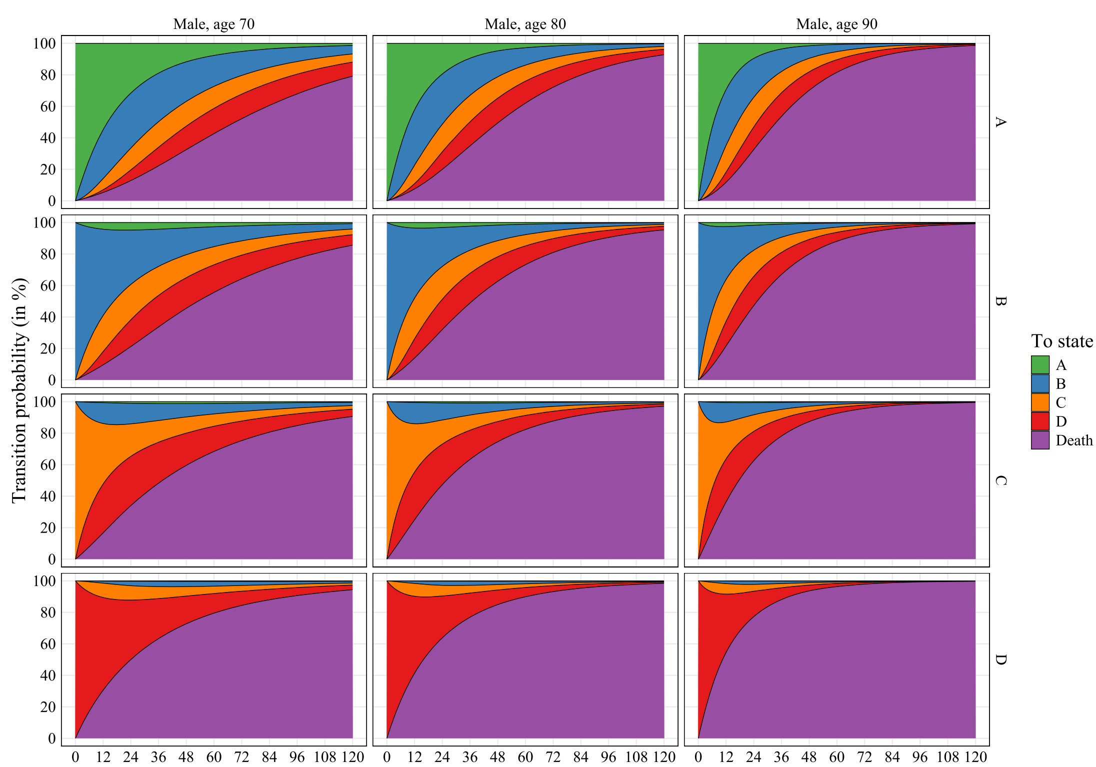
To study the effects of levels of dependence and medical diagnoses, we select four other health profiles (see Section 3.1) that offer the most significant insights or present unique characteristics. In the presentation of the results, we focus on the transition probabilities and associated costs for an 80-year-old individual, as shown in Figs. 9 and 10. Specifically, we analyze the second and third most common profiles: general severe conditions (short: severe) and moderate-severe conditions with nervous diseases (short: nervous); the fifth most common profile: moderate conditions with cerebrovascular diseases (short: cerebrovascular); and the eighth profile, noted for the shortest duration of stay: moderate conditions with tumor diseases (short: tumor). We omit the remaining three profiles, characterized by moderate conditions with endocrine, respiratory, and blood diseases from the detailed discussion. Indeed, these pathology groups are not explicitly included in our multi-state model, which limits the extent to which their specific impacts can be assessed.

General severe conditions. This profile is characterized by having six diagnoses ($ND = 6$), with median values of 9, 6, and 13 for the mental, heart, and other pathology scores, respectively. The levels of dependence are the highest among all eight groups with $DP = 8$, $PM = 8$, and $OR = 6$. Finally, the visual and hearing impairments are classified as “mild”.

The transition probabilities for the severe profile in Fig. 9 (see the short-hand notation “severe”) present distinct patterns compared to the baseline health profile (see Fig. 7). Individuals transition to state D more rapidly and maintain a higher probability of staying in state D throughout the observed period. Both male and female individuals in the severe profile exhibit a higher and earlier transition to death, reflecting

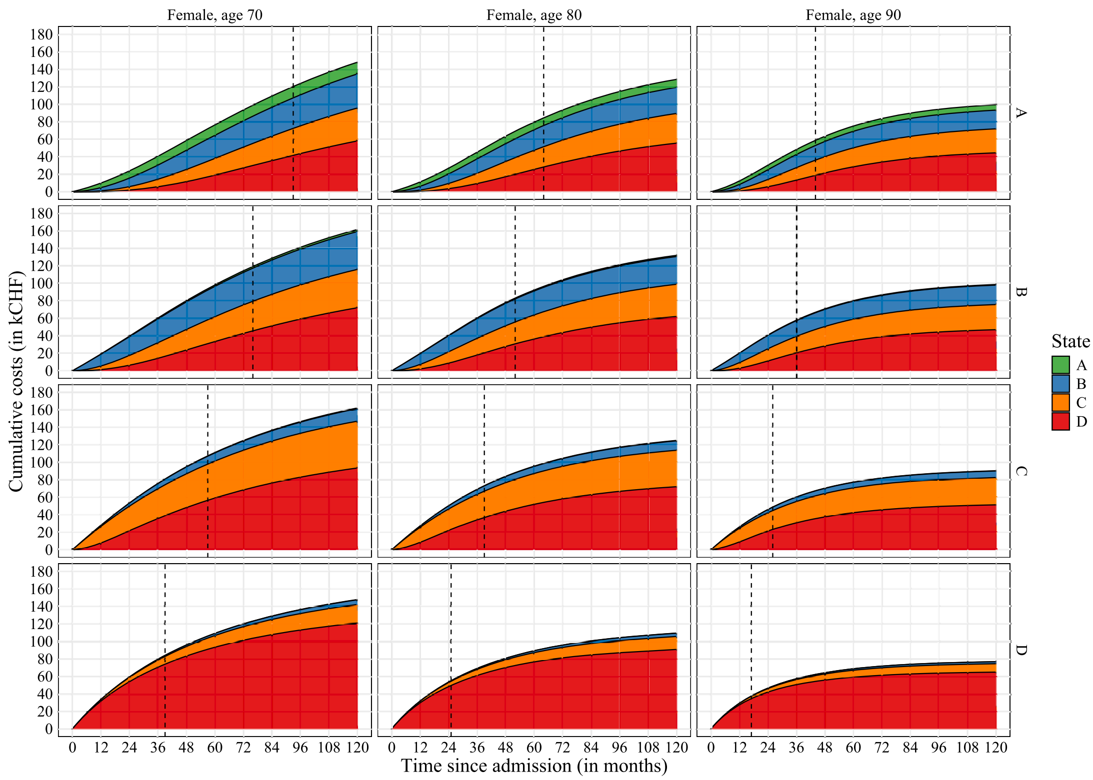


(a) Females.

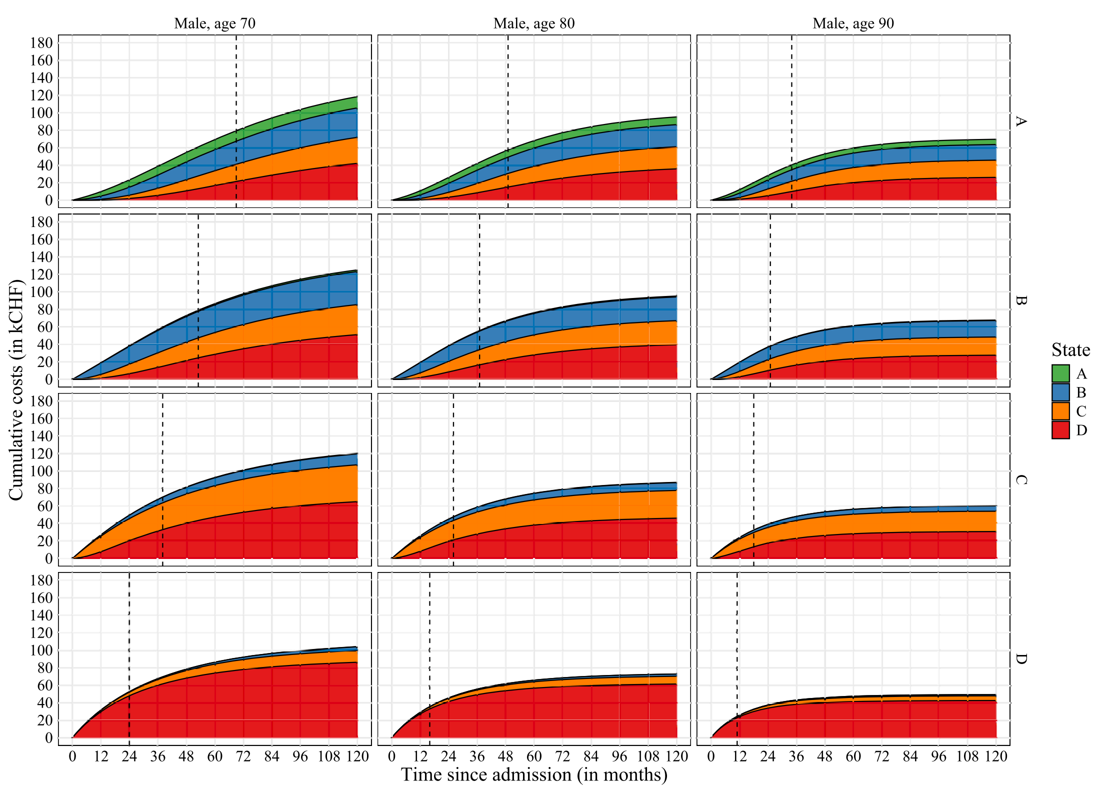


(b) Males.

Fig. 7. Transition probabilities for females and males in the baseline health profile.



(a) Females.



(b) Males.

Note: Vertical dashed lines indicate the median survival times.

Fig. 8. Average cumulative costs for females and males in the baseline health profile.

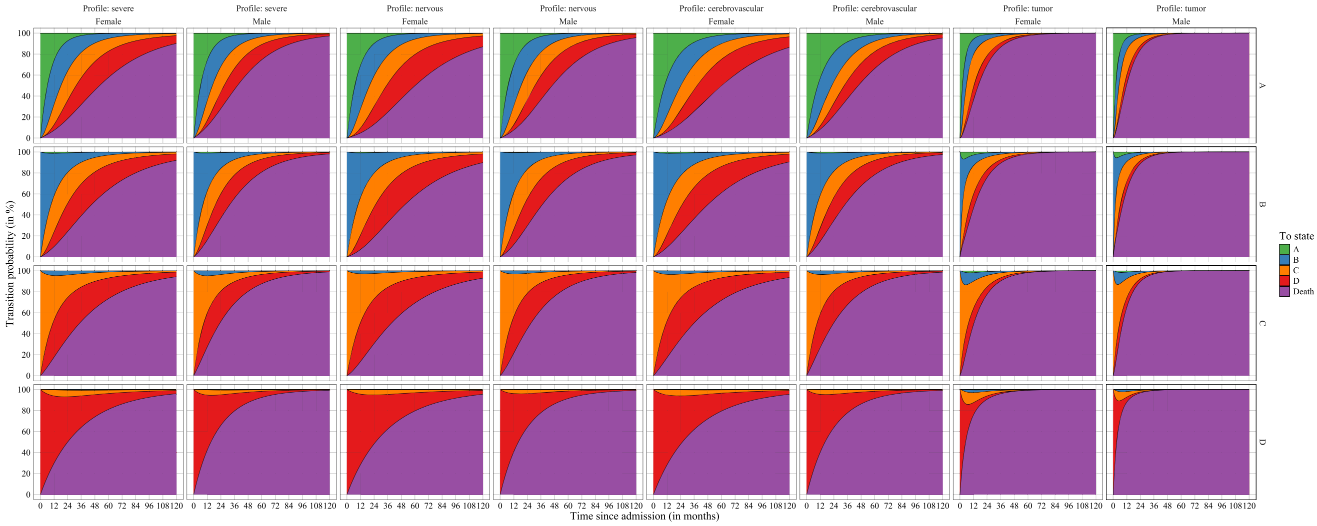
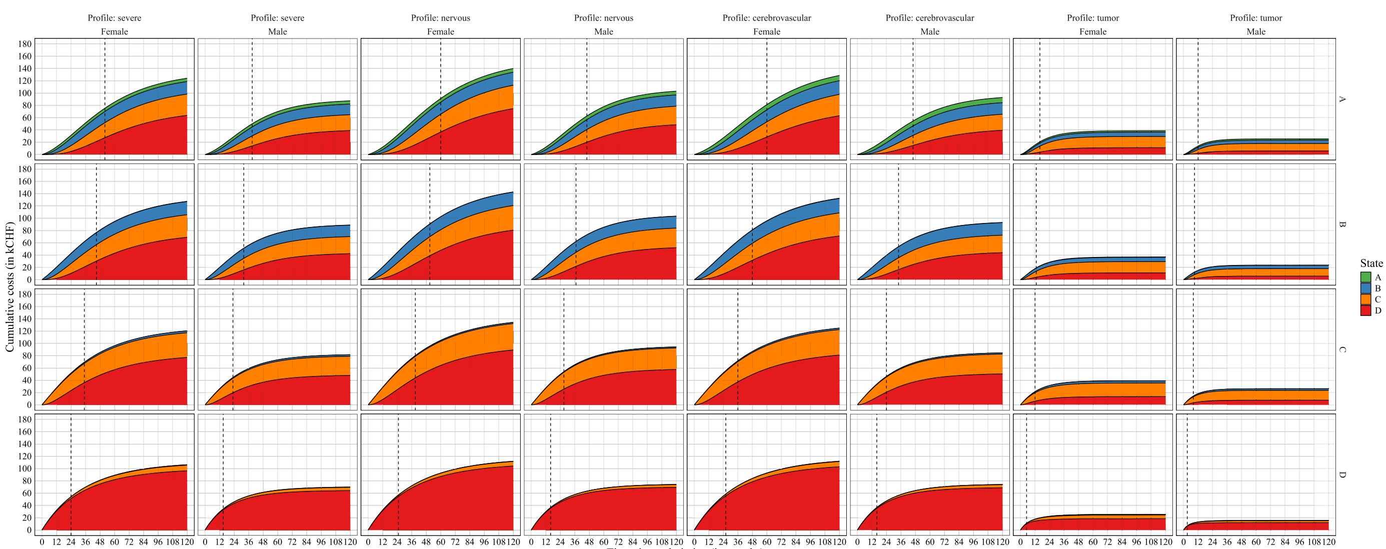


Fig. 9. Transition probabilities for an 80-year-old elderly person in selected health profiles.



Note: See Fig. 8.

Fig. 10. Average costs for an 80-year-old elderly person in selected health profiles.

the greater health burden and higher levels of dependence. Moreover, the transition from states A and B to higher states occurs quicker in the severe profile than in the baseline. This reflects the significant impact of more severe medical conditions and higher dependence levels compared to the generally healthier baseline group. From Fig. 10, we observe that females' costs rise more steeply initially, especially from states C and D, reflecting higher early dependency, while males' costs increase more gradually. This pattern indicates that in severe conditions, while females rapidly transition to high-dependency states, implying higher costs, males progress into these states at a slower rate. However, both genders eventually accumulate comparable costs by the end of the observed period compared to the baseline profile, highlighting the significant financial impact of high-dependency care over time.

Studies provide support for the connections described here related to severe conditions in LTC, transition probabilities, and associated costs. Esquível et al. (2021) introduce a non-homogeneous continuous-time Markov chain model for LTC, considering different dependence levels and states of care, including severe conditions. The study shows how individuals with higher dependence levels transition more rapidly to severe states, increasing the overall cost burden in LTC, which matches our observations. Similarly, Deshmukh (2012) focus on the financial impact of transitioning between care states in LTC settings, exploring how severe conditions lead to higher costs over time due to extended stays in higher dependency states. Finally, Hill et al. (2021) emphasize the importance of accounting for varying transition rates in multi-state models, highlighting that severe conditions can drastically alter transition probabilities and cost projections.

Moderate-severe conditions with nervous diseases. This health profile is the third most prevalent and similar to the severe profile, yet it is distinguished by a notable prevalence of medical diagnoses from the nervous group. In particular, the median number of diagnoses is smaller by one ($ND = 5$), followed by a redistribution among the pathology scores (mental 8, heart 4, and other = 10), complemented by the median score of 9 in “nervous” pathologies. The levels of dependence and visual and hearing impairments are identical to those observed in the severe profile, with the exception of physical mobility, which is reduced by one unit ($PM = 7$).

The nervous profile (short “nervous” in Figs. 9 and 10) exhibits longer median survival times across all states in comparison to the severe profile. Upon initial admission in state D, transition probabilities to the lower state C are notably less frequent in the nervous profile, which contributes to prolonged stays in state D. This indicates a sustained higher level of dependency and increased care costs. Furthermore, the similar levels of dependence and impairment to those observed in the severe profile highlight the nuanced differences in care needs driven by nervous pathologies. The cumulative cost curves for the nervous and severe profiles demonstrate comparable behaviors during the initial months, indicating comparable initial care costs across the profiles. However, as time progresses, the nervous profile incurs higher cumulative costs, especially in state D, where prolonged high-level care leads to increased expenditure. This sustained higher cost in state D aligns with the longer stay observed in the transition probability analysis, emphasizing the financial implications of managing chronic nervous conditions in LTC settings. In summary, while the initial cost impact is similar between the profiles, the long-term financial burden is greater for the nervous profile due to extended periods of high-dependency care.

These findings align with the literature indicating that nervous conditions contribute to prolonged dependency and increased costs in LTC settings. Nihtilä et al. (2007) discuss the increased risk of institutionalization and higher costs for nervous system conditions. Turner-Stokes et al. (2008) emphasize the complexity of managing long-term neurological conditions and their associated costs, while Brandis and Stacom (2009) provide insights into the high levels of dependency and care costs in patients with multiple sclerosis, reflecting the observed patterns for the nervous health profile.

Moderate conditions with cerebrovascular diseases. Ranking as the fifth most prevalent, this profile typically encompasses individuals diagnosed with six medical conditions ($ND = 6$). The median pathology scores for these patients are 7, 6, and 9 for the mental, heart, and other scores, respectively, and notably 8 for the cerebrovascular score. The levels of dependence and visual and hearing impairments are one unit higher than those observed in the baseline group and mirror those observed in the nervous profile. In comparison, the cerebrovascular profile has a lower median score for other pathologies at 9 against 13 in the severe profile, similar levels of mental and heart conditions, and a one-unit lower score in physical mobility.

The transition probability graphs for the cerebrovascular profile (short: cerebrovascular) in Fig. 9 show a longer survival time compared to the severe profile, with individuals tending to remain longer in their initial state, particularly in states A and B. For both females and males in the cerebrovascular profile, transitions to higher dependency states and death take place later in time (broader curves), suggesting a slower progression of care needs. For females diagnosed with cerebrovascular conditions, the cost trajectories consistently accumulate higher costs over time compared to those with severe profiles. Notably, when starting from states B, C, or D, the costs align more closely with those observed in the nervous profile, indicating a substantial financial burden. Conversely, for males diagnosed with cerebrovascular conditions, the cost trajectories for those starting in states A and B reach slightly higher levels after 10 years compared to those with severe profiles, suggesting a slight increase in LTC costs. However, males starting in states C and D exhibit almost identical cost trajectories across both profiles, indicating that severe conditions and cerebrovascular diseases impose a comparable financial burden in these states. Overall, while cerebrovascular diseases tend to increase the cumulative costs of LTC, particularly for females, the impact on males is less pronounced and varies significantly based on the initial state of care.

These observations align with studies on cerebrovascular diseases in LTC settings. Tamiya et al. (2001) show that cerebrovascular diseases lead to increased care needs over time, supporting the pattern of prolonged stays and higher care costs observed in our analysis. Similarly, Xu et al. (2018) demonstrate that the burden of cerebrovascular disease predicts functional deterioration and increased LTC costs, consistent with the significant financial burden identified for the cerebrovascular profile. Additionally, Kalaria (2012) highlight the link between cerebrovascular disease and cognitive impairment, contributing to greater care needs and longer stays in institutional settings, which reflects our findings for this profile.

Moderate conditions with tumor diseases. This group, identified as the least prevalent, is distinguished by a higher number of medical diagnoses ($ND = 8$). Individuals in this category show median pathology scores of 4, 9, and 17 for heart, other, and tumor, respectively. Relative to the baseline profile, this group shares similar levels of dependence and sensory impairments but has slightly increased physical mobility by one level.

The tumor profile (short: tumor) demonstrates a remarkably higher mortality rate compared to other profiles. Particularly for those starting in state C, there is a prolonged period of stability before an eventual shift to the death state, indicating sustained intensive care needs. An unusual pattern emerges in states A and B, where individuals are more likely to remain in state A or revert to it within the first year of admission, unlike other profiles, which typically show a progression to higher states. This evolution is specific to individuals with tumor diseases. The cost graphs for the tumor profile show relatively smaller costs, as median survival times do not surpass twelve months, particularly in states C and D. Compared to other health profiles, the tumor profile distinctly features rapid transitions to death and a higher likelihood of remaining in or returning to the lowest care states. These observations underscore the tumor profile’s distinct impact on LTC costs, where high initial care needs are offset by significantly reduced life expectancy.

These findings align with the literature on cancer care in LTC settings. Burke et al. (2023) discuss the challenges of managing advanced cancer within LTC facilities, highlighting the rapid progression of disease and intensive care coordination needs, which reflect the rapid transitions observed in our tumor profile. Similarly, Lin (2017) discusses how cancer patients often require intensive initial care, but reduced life expectancy leads to distinct cost patterns, consistent with our findings. Lastly, Jacobs (2012) describe the long-term and late effects of cancer treatments, which often result in complex care needs and aggressive progression, corroborating the unique care trajectories observed for the tumor profile in LTC.

4.5. Summary and discussion of the results

Overview of results. Our study of institutionalized elderly across various health profiles has yielded detailed insights into the transitions and associated LTC costs. Notably, the model demonstrated a robust fit for the majority of conditions. However, it also indicated a potential poor estimation in states A, D, and death, which could affect the accuracy of transition and survival predictions.

Key findings emerged from analyzing different profiles: the baseline health profile, most prevalent among the elderly, indicated that females generally incur higher LTC costs due to their prolonged care needs. The two profiles with general severe conditions and moderate-severe conditions with nervous diseases highlighted rapid progression to high-dependency states with considerable initial costs, particularly prominent among females who transitioned quicker to these states. Females with cerebrovascular conditions often experience slower progression to higher dependency states but eventually accumulate higher costs, suggesting that strategies specific to medical conditions might be necessary to manage care effectively. The profile characterized by tumor diseases profile was particularly notable for its rapid transitions to death, resulting in lower overall costs due to shorter survival times, presenting a unique economic dynamic compared to other conditions.

These insights highlight the critical need for precise model fitting and the development of care strategies that account for age, gender, and specific health conditions. This understanding is crucial for policymakers and healthcare providers to optimize resource allocation and improve care outcomes for the aging population. It also underscores the necessity of strategic interventions in managing severe conditions to alleviate their financial impacts.

Table 4 provides a breakdown of the average time $E_{rs}(\delta, \mathbf{z})$ (see Equation (6)) an individual starting in state r is expected to spend in the care state s, stratified by health profile determined by $\mathbf{z}$ including gender and age, up to their median survival time δ, excluding any duration spent in the state of death. Specifically, we present results showing the average time spent in each care state E_{rA}, E_{rB}, E_{rC} and E_{rD} and the total expected costs C_r (see Equation (7)) up to the point where 50% of individuals are expected to have passed away. Aligning with previous analysis, we detail the results for individuals admitted at the age of 80. Summary data for 70 and 90-year-old admitted individuals are included as complementary age groups to offer insights into how care needs and associated costs vary with age. The "Prevalence" column reflects the distribution of individuals within each profile and gender in the overall dataset, providing information on the typicality of each scenario. This table is pivotal for understanding the care needs and financial implications associated with different health profiles in institutional LTC, aiding in strategic planning and resource allocation to efficiently meet the needs.

Nursing resources. The columns $E_{rs}(\delta)$ with $s = \mathrm{A, B, C, D}$ in Table 4 provide relevant metrics that directly impact nursing requirements in LTC settings. The numbers provide the average duration elderly individuals spend in the four care states before transitioning to death. Analyzing the length of stay for various health profiles reveals distinct patterns in managing care needs across both genders and the initial care states. For instance, in the baseline health profile, females admitted to the institution in state A experience longer durations across all states compared to males (e.g., 14.9 months for females and 14 for males in state A), which indicates a prolonged need for lighter care levels. In contrast, in the severe and nervous profiles, both genders exhibit shorter stays in lower states like A and B but consistently spend more time in the highest dependency state when starting in state D. For example, females in the nervous profile, on average, spend $E_{\mathrm{DD}}(\delta) = 17$ months in state D, while males spend 11 months. In comparison, females and males in the healthy profile spend 15.5 and 10.2 months, respectively, in state D. Notably, individuals with cerebrovascular conditions exhibit a similar pattern to the baseline health profile. In particular, they have slightly lower lengths of stay in the lower care states and slightly longer durations in the higher-intensity states. Finally, the tumor-afflicted individuals, regardless of gender, exhibit significantly reduced E_{rs} across all states due to accelerated deterioration of health, with a distinctive tendency to spend the majority of their time in the state they entered the institutional LTC.

This complex nature of LTC demands a nuanced approach to nursing, especially in managing prolonged care in higher dependency states. The data highlights the extended periods in states C and D for conditions like cerebrovascular and nervous diseases, where patients often require intense and sustained care. This situation is particularly critical for females who demonstrate a need for prolonged high-level care, underscoring the importance of gender-specific care strategies and resource allocation. To effectively address these diverse and complex care requirements, a well-trained nursing workforce is essential. Continuous education and specialized training are crucial to equip caregivers with the skills necessary for managing these complex health profiles. Furthermore, the Swiss healthcare system is constrained by a deficit of qualified local caregivers (Zúñiga et al., 2010; Haller et al., 2015), which reflects the necessity for supportive immigration policies that facilitate the influx of competent care providers (Nichols et al., 2010). These strategies are essential for maintaining high standards of care, improving patient outcomes, and adequately responding to the evolving needs of an aging population in institutional LTC settings. By investing in educational advancement and incorporating a strategy that includes gender and profile-specific care planning, LTC facilities can optimize staffing and resource use, ensuring that the aging population's dynamic demands are met effectively.

Infrastructure. The median survival times offer a good perspective on infrastructure needs in LTC settings, revealing how long individuals are expected to utilize care facilities. This data is essential for planning future care infrastructure and resource allocation within these institutions. Across all health profiles and genders, the median survival times decrease significantly with increasing age at entry. For example, females in the baseline health profile starting in state A exhibit a median survival time of 93 months at age 70, which drops to 44 months at age 90. This illustrates a marked decline in longevity as age increases, reflecting greater immediacy in care needs and infrastructure planning for older entrants. Focusing on $AG = 80$, females typically demonstrate longer survival times across all profiles and starting states, which is particularly pronounced in the baseline health, nervous, and cerebrovascular profiles. For instance, baseline health females starting in state A have a median survival time of $\delta = 64$ months, compared to 49 months for their male counterparts. Similarly, females in the cerebrovascular profile starting in state A have a median survival time of 60 months versus 46 months for males, indicating a substantial gender disparity in care duration that could impact resource planning. This trend persists in the severe profile, showing smaller δ, while maintaining the gender disparity. For example, females in the severe health profile starting in state D have a median survival time of 24 months, compared to 15 months for males. Tumor profiles present the most drastic differences, with extremely short durations, highlighting a distinct infrastructure challenge. Tumor-affected females starting in state D have a survival time of only 5 months at age 80, significantly lower than other profiles. This disparity underscores the necessity for LTC facilities to adapt their infrastructure to accommodate not only the varying lengths of stay associated with different medical conditions, coming with a distinct prevalence but also the specific needs that arise from gender differences in survival rates.

The observed variations in median survival times have direct implications for LTC infrastructure planning. Facilities must ensure they have sufficient beds and appropriately configured rooms to accommodate the different types and durations of stay that can be anticipated for each health profile. For instance, the significantly shorter median survival times for older entrants across all profiles, such as tumor patients at age 90 having median times as low as 2 months, suggest a need for flexible room allocations that can adapt to high turnover rates. Conversely, profiles with longer survival times, such as baseline health females entering at age 70 with survival times up to 93 months, require stable, long-term accommodations. Additionally, the data indicates a potential shift in care strategy, where individuals with longer predicted survival times and less intensive care needs, such as those in the baseline health profile, could benefit from expanded home-care services. This shift could alleviate pressure on LTC institutions by reducing the demand for in-facility resources, allowing these institutions to focus on patients with more severe conditions who require intensive, specialized care. These strategic infrastructure adjustments are crucial for optimizing care delivery, infrastructure, and resource allocation in response to the aging population's diverse needs.

Basic health insurance. The analysis of expected costs at the median survival times provides insights into the financial implications of different health profiles on nursing costs covered by basic health insurance in Switzerland. Notably, females incur higher costs compared to males, reflecting longer survival times and potentially more intensive care needs. For instance, nervous conditions in females aged 80 show an average cost of $C_{\mathrm{A}}(\delta) = 91.2$ thousand Swiss francs, significantly higher than their male counterparts at 62.7 thousand. This trend persists across profiles and ages, with younger individuals ($AG = 70$)

Table 4
Time spent in care and care costs up to the median survival time by health profile, gender, and age.

			$AG = 80$						$AG = 70$		$AG = 90$	
Gender	Initial state r	Preva-lence	$E_{r\mathrm{A}}(\delta)$	$E_{r\mathrm{B}}(\delta)$	$E_{r\mathrm{C}}(\delta)$	$E_{r\mathrm{D}}(\delta)$	δ	$C_r(\delta)$	δ	$C_r(\delta)$	δ_r	$C_r(\delta)$
Baseline health profile (26.3% of the data)												
Female	A	(29.7)	14.9	16.6	9.9	8.8	64	84.3	93	120.3	44	58.7
	B	(46.0)	1.7	17.9	10.8	9.5	52	82.9	76	119.2	36	57.8
	C	(18.0)	0.3	4.1	12.9	11.5	39	73.2	57	107.4	26	48.9
	D	(6.4)	0.0	0.4	2.0	15.5	25	55.1	39	84.1	17	37.6
Male	A	(25.9)	14.0	12.9	6.6	4.7	49	57.3	69	79.3	34	40.3
	B	(47.6)	1.1	14.5	7.4	5.1	37	55.7	53	78.8	25	38.1
	C	(18.5)	0.1	2.9	9.6	6.6	26	47.7	38	70.2	18	32.6
	D	(8.1)	0.0	0.2	1.0	10.2	16	35.5	24	52.8	11	24.3
Profile: severe (19.0% of the data)												
Female	A	(0.5)	9.1	12.5	10.4	8.5	52	75.3	75	107.6	36	52.3
	B	(13.7)	0.4	13.4	11.4	9.4	45	76.8	66	112.1	31	52.8
	C	(37.3)	0.0	1.3	13.4	11.2	35	69.3	50	100.8	24	46.9
	D	(48.5)	0.0	0.1	1.2	16.0	24	54.2	36	80.8	16	36.4
Male	A	(0.2)	8.8	10.0	7.0	4.4	39	50.0	55	69.8	27	34.9
	B	(10.4)	0.3	11.2	8.0	5.0	32	51.2	46	73.5	22	35.2
	C	(34.5)	0.0	0.9	10.0	6.1	23	44.5	34	66.5	16	30.2
	D	(54.9)	0.0	0.0	0.6	10.2	15	34.2	23	51.8	10	23.0
Profile: nervous (18.2% of the data)												
Female	A	(3.2)	10.3	13.4	12.4	11.4	60	91.2	86	129.6	42	64.1
	B	(18.9)	0.3	14.1	13.2	12.1	51	90.8	74	131.2	36	63.8
	C	(37.1)	0.0	0.9	14.9	13.7	39	80.2	56	116.3	27	54.9
	D	(40.8)	0.0	0.0	0.9	17.0	25	57.0	38	85.8	17	38.7
Male	A	(1.8)	10.1	11.1	8.8	6.2	46	62.7	65	87.6	32	43.9
	B	(14.8)	0.2	12.1	9.7	6.8	37	62.2	53	88.8	26	43.7
	C	(29.8)	0.0	0.7	11.7	8.0	27	54.1	39	78.9	19	37.4
	D	(53.6)	0.0	0.0	0.4	11.0	16	36.6	24	54.5	11	25.0
Profile: cerebrovascular (12.8% of the data)												
Female	A	(6.8)	14.3	13.4	10.5	8.9	60	81.2	86	115.4	42	57.0
	B	(20.8)	0.5	14.8	11.8	9.8	48	81.1	70	117.9	33	55.6
	C	(27.5)	0.0	1.0	14.0	11.6	36	71.7	52	105.3	24	47.4
	D	(44.9)	0.0	0.1	1.2	17.4	26	58.6	38	85.7	17	38.8
Male	A	(5.3)	13.7	10.6	7.0	4.6	46	54.5	65	76.2	32	38.1
	B	(17.1)	0.3	12.3	8.2	5.1	34	53.8	49	77.4	23	36.5
	C	(23.9)	0.0	0.7	10.6	6.4	24	46.4	35	68.8	16	30.5
	D	(53.7)	0.0	0.0	0.5	11.0	16	36.6	24	54.5	11	25.1
Profile: tumor (1.1% of the data)												
Female	A	(8.2)	3.6	3.5	4.1	1.2	16	20.7	23	28.8	11	14.7
	B	(30.3)	0.6	3.5	4.2	1.2	13	19.2	18	25.9	9	13.7
	C	(26.2)	0.2	1.2	5.9	1.9	12	21.5	17	30.5	9	15.8
	D	(35.2)	0.0	0.0	0.4	3.1	5	11.0	8	17.0	3	6.9
Male	A	(3.6)	3.3	2.6	2.6	0.6	12	13.8	16	17.9	8	9.5
	B	(30.4)	0.4	2.7	2.7	0.6	9	12.4	12	16.2	6	8.6
	C	(27.7)	0.1	0.8	4.2	1.0	8	14.3	12	21.0	6	10.5
	D	(38.4)	0.0	0.0	0.2	2.0	3	6.8	5	10.8	2	4.5

Notes: The column "Prevalence" indicates the distribution in % per profile and per gender. The median survival time δ indicates the time in months where the probability of death reaches 50%. The average times of stay E_{rs} and the total costs C_r are calculated with regard to the time δ and expressed in months and kCHF, respectively.

incurring higher costs due to longer survival periods. Comparatively, tumor profiles exhibit much lower costs across all ages due to significantly shorter survival times, emphasizing the rapid progression to death. For example, tumor-affected females aged 80 have costs of $C_{\mathrm{A}}(\delta) = 20.7$ thousand, which is considerably lower than those with cerebrovascular conditions at 81.2 thousand. Across various health profiles, a clear pattern emerges where the costs associated with initial higher care states (such as states C and D) tend to be lower compared to those starting from lower states like A and B, particularly for those entering at older ages ($AG = 90$). This trend is largely attributed to increased mortality rates in higher initial states, shortening the duration of care and thus reducing cumulative costs. However, for individuals entering at a younger age ($AG = 70$), this pattern shifts notably for severe, nervous, and cerebrovascular profiles, where the highest costs are often recorded for those starting in state B, suggesting prolonged care durations before reaching higher mortality states. In contrast, the tumor profile uniquely shows the highest costs from state C, indicating specific care dynamics associated with this condition.

Our results provide insights for basic health insurers and policymakers in efficiently planning and allocating resources to meet the diverse needs of the aging population. By understanding the expected costs linked to various health profiles and entry states, policies and pricing models can be refined to reflect the true financial risk associated with different levels of care. Additionally, this data allows policymakers to better forecast LTC funding requirements and develop strategies to ensure that essential care services are sustainable and accessible. Such detailed cost analysis aids in the financial planning of public health services, ensuring that funds are utilized effectively. Furthermore, it can support the optimization of private insurance packages.

Private insurance. Private insurance plays a pivotal role in supplementing the shortcomings of basic health insurance, particularly in the coverage of out-of-pocket expenses, including lodging, meals, and specialized medications that are not reimbursed under social insurance policies. The median survival times, δ, derived from our model, provide crucial insights for private insurers, as they can use these durations to estimate costs associated with per diems, lodging, and meals over the expected period an individual will require LTC. This approach allows insurers to assess the premiums required upfront to cover these ongoing costs effectively. Additionally, the average lengths of stay in each care state, E_{rs}, facilitate the development of personalized insurance products tailored to the intensity of care an individual is likely to require. This personalized approach not only ensures that individuals receive the appropriate level of support and care but also helps insurers manage risks and resources more effectively. In the case of LTC insurance products with a savings component proposed to individuals before they require any care, the prevalence of different health profiles by gender highlighted in Table 4 provides insurers basic insight for weighting different levels of care demand. In addition, our results enable, for example, the pricing of insurance products that can be made available to elderly individuals at the moment they are admitted to an institution. Using the age, gender, and health care profile of a person at entry, the insurer could offer to cover the expected out-of-pocket expenses until death against a lump-sum payment. We believe that our approach enhances the base of knowledge for private insurers to provide robust financial solutions that support individuals throughout their time in LTC, ensuring that all necessary expenses are covered comprehensively.

5. Conclusions

In this study, we conducted a comprehensive analysis of a private dataset from nursing homes in the Canton of Geneva, Switzerland, encompassing 21 494 elderly individuals aged 65 or older. Our research utilized a multi-state Markov model to assess transitions between grouped care states – ranging from quasi-autonomy to severe dependency – within the Swiss social health insurance framework. By systematically grouping the care levels and focusing on significant variables at admission, such as demographic information, medical diagnoses, and levels of dependence, we have identified key patterns and trends in the evolution of care needs over time. This approach not only facilitated a clearer understanding of the longitudinal care dynamics, but also allowed us to model the long-term costs associated with different levels of care required by the elderly in institutional LTC settings.

We aggregated the twelve care levels of the Swiss system into four broader categories, ranging from minimal assistance to severe dependency. This classification enables comparison with other studies and provides a clear framework for assessing the impact of various health conditions on LTC trajectories. Utilizing common health profiles among institutionalized elderly allowed us to analyze the influence of demographic and medical covariates on transition probabilities and associated costs. The baseline health profile, which is the most prevalent, incurs higher LTC costs due to extended care durations. In contrast, profiles characterized by severe conditions and nervous diseases show a rapid progression to higher dependency states, resulting in considerable initial costs. Particularly, females in these profiles transition more quickly to high-dependency states, highlighting the need for targeted care strategies that consider both medical and demographic factors. Individuals with cerebrovascular conditions tend to have a slower progression to higher care states but eventually accumulate higher costs, suggesting that prolonged care interventions are necessary. Conversely, the tumor profile is marked by rapid transitions to death, resulting in lower overall costs due to shorter survival times. These distinct patterns emphasize the importance of adapting care and financial planning to the specific health profiles of elderly individuals in institutional LTC settings.

The integration of demographic information, medical diagnoses, and levels of dependence at the point of admission enables the model to provide insights for the planning of care and infrastructure, as well as the design of insurance products, addressing both public and private sectors. Analysis of health profiles revealed a nuanced variation in care needs, depending on the initial state of both genders at different ages. This variation in care duration underscores the necessity for advanced nursing strategies and gender-specific care planning. Additionally, the data on median survival times is crucial for predicting infrastructure needs, indicating a requirement for facilities to adapt to varying lengths of stay and high turnover rates, particularly for older entrants and patients with rapidly progressing conditions like tumors. The analysis relevant to basic health insurance demonstrates that costs are influenced by the patient's health profile and age at entry, with women generally incurring higher costs due to longer survival times. Conversely, those in more intensive initial care states accumulate lower costs due to decreased care durations. Private insurers can utilize the insights into projected median survival times and expected care state durations to develop insurance products that accurately reflect the costs and care needs of LTC patients. This enables the provision of comprehensive solutions that meet all necessary expenses not covered by the basic insurance scheme.

While our study offers valuable insights through the use of a multistate Markov model, it lacks a comparison between the estimated costs derived from the model and the actual costs documented in real data. Such a comparison could enhance the validity of our findings by aligning the obtained predictions with practical outcomes. Additionally, our analysis is constrained by the nature of our panel data, which captures the health states only at discrete intervals. This limitation prevents us from precisely determining the exact times of transitions between care states, leading to potential discrepancies in the fit of our model. Furthermore, our decision to aggregate the twelve care levels into four broader categories, though based on practical and clinical considerations to simplify the analysis and enhance interpretability, may contribute to these discrepancies. While each category corresponds to intuitive ranges of care hours per day, this grouping might oversimplify the underlying transition dynamics, potentially masking subtle differences in transition probabilities between individual care levels. Moreover, the study's reliance on fixed covariates at the point of admission restricts our ability to account for changes in an individual's condition over time, potentially skewing the assessment of transition probabilities and cost implications. Assuming the transitions occur at the time of observations, a semi-Markov approach that allows the transition probabilities to depend not only on the current state but also on the duration of stay in that state and changing covariates could potentially address this issue.

Future research could significantly benefit from incorporating time-varying covariates into the models used to predict transitions and costs in institutional LTC. Allowing variables such as health status, level of dependence, and medical conditions to change over time, would offer more accurate predictions. Additionally, exploring joint modeling approaches where the evolution of care intensity directly influences survival probabilities could provide a more dynamic understanding of the development of LTC needs. Such models could uncover the interdependencies between care requirements and survival, leading to more effective care planning and potentially improving patient outcomes by allowing for more personalized and timely interventions in care strategies.

CRediT authorship contribution statement

Aleksandr Shemendyuk: Writing – original draft, Visualization, Validation, Software, Resources, Methodology, Investigation, Formal analysis, Data curation. **Joël Wagner:** Writing – review & editing, Supervision, Conceptualization.

Declaration of competing interest

None declared.

Acknowledgements

We are thankful to the General Directorate for Health of the Republic and Canton of Geneva for providing the data used in this study.

Appendix A

Table 5
Prevalence of states at different times across selected covariates derived from Aalen-Johansen estimator.

Variable	State	Entry			After 1 year			After 5 years			After 10 years		
		M	%	$\hat{p}$	M	%	$\hat{p}$	M	%	$\hat{p}$	M	%	$\hat{p}$
Age at entry AG													
65-69	A	80	0.4	12.7	73	0.3	11.6	32	0.1	6.0	15	0.1	3.5
	B	203	0.9	32.3	174	0.8	28.6	70	0.3	13.9	30	0.1	8.3
	C	144	0.7	22.9	139	0.6	22.9	63	0.3	13.7	23	0.1	5.9
	D	201	0.9	32.0	152	0.7	25.1	105	0.5	21.2	38	0.2	11.0
70-79	A	457	2.1	11.9	394	1.8	10.6	143	0.7	4.2	42	0.2	1.3
	B	989	4.6	25.8	793	3.7	21.6	228	1.1	7.2	59	0.3	2.1
	C	1022	4.8	26.7	865	4.0	23.6	280	1.3	9.3	77	0.4	3.0
	D	1358	6.3	35.5	1115	5.2	30.2	667	3.1	22.0	206	1.0	8.7
80-89	A	1399	6.5	12.5	1196	5.6	10.9	287	1.3	2.9	39	0.2	0.4
	B	3112	14.5	27.8	2438	11.3	22.6	497	2.3	5.5	70	0.3	0.9
	C	3184	14.8	28.4	2604	12.1	24.5	675	3.1	7.6	88	0.4	1.2
	D	3505	16.3	31.3	2668	12.4	25.0	1416	6.6	16.3	266	1.2	3.8
90-99	A	527	2.5	9.3	413	1.9	7.5	38	0.2	0.8	2	0.0	0.0
	B	1602	7.5	28.2	1162	5.4	21.6	127	0.6	2.9	9	0.0	0.2
	C	1640	7.6	28.9	1179	5.5	22.1	188	0.9	4.4	6	0.0	0.2
	D	1915	8.9	33.7	1310	6.1	24.4	402	1.9	9.5	25	0.1	0.8
100+	A	3	0.0	1.9	2	0.0	1.9	0	0.0	0.0	0	0.0	0.0
	B	17	0.1	10.9	14	0.1	9.0	1	0.0	0.7	0	0.0	0.0
	C	46	0.2	29.5	23	0.1	15.7	0	0.0	0.0	0	0.0	0.0
	D	90	0.4	57.7	50	0.2	33.3	4	0.0	3.1	0	0.0	0.0
Gender GE													
Female	A	1905	8.9	12.2	1623	7.6	10.6	394	1.8	2.8	76	0.4	0.6
	B	4446	20.7	28.5	3478	16.2	23.3	737	3.4	5.8	140	0.7	1.3
	C	4490	20.9	28.8	3676	17.1	24.8	1014	4.7	8.3	167	0.8	1.6
	D	4747	22.1	30.5	3798	17.7	25.4	2163	10.1	17.8	465	2.2	4.9
Male	A	561	2.6	9.5	455	2.1	7.9	106	0.5	2.1	22	0.1	0.5
	B	1477	6.9	25.0	1103	5.1	19.5	186	0.9	4.0	28	0.1	0.8
	C	1546	7.2	26.2	1134	5.3	20.5	192	0.9	4.4	27	0.1	0.8
	D	2322	10.8	39.3	1497	7.0	26.8	431	2.0	9.9	70	0.3	2.2
Primary diagnosis $D1$													
Mental	A	763	3.5	10.3	632	2.9	8.7	149	0.7	2.3	29	0.1	0.5
	B	2104	9.8	28.4	1669	7.8	23.6	341	1.6	5.8	71	0.3	1.4
	C	2212	10.3	29.9	1822	8.5	25.9	483	2.2	8.6	77	0.4	1.8
	D	2324	10.8	31.4	1843	8.6	26.1	968	4.5	17.4	220	1.0	5.2
Cerebro-vascular	A	75	0.3	6.0	61	0.3	5.0	24	0.1	2.2	2	0.0	0.2
	B	221	1.0	17.8	171	0.8	14.6	26	0.1	3.0	8	0.0	0.9
	C	305	1.4	24.5	246	1.1	20.5	58	0.3	5.8	14	0.1	1.6
	D	644	3.0	51.7	470	2.2	39.6	184	0.9	18.9	42	0.2	5.6
Respi-ratory	A	49	0.2	14.4	35	0.2	10.6	7	0.0	2.3	0	0.0	0.0
	B	103	0.5	30.2	77	0.4	23.8	11	0.1	3.7	1	0.0	0.4
	C	81	0.4	23.8	54	0.3	16.6	13	0.1	4.9	1	0.0	0.4
	D	108	0.5	31.7	63	0.3	19.0	15	0.1	5.4	4	0.0	1.6
Blood	A	12	0.1	14.5	10	0.0	12.0	3	0.0	4.4	0	0.0	0.0
	B	27	0.1	32.5	18	0.1	21.7	0	0.0	1.6	0	0.0	0.0
	C	21	0.1	25.3	15	0.1	18.9	1	0.0	1.6	0	0.0	0.0
	D	23	0.1	27.7	14	0.1	17.0	6	0.0	9.4	1	0.0	1.7
Nervous	A	152	0.7	3.9	112	0.5	2.9	20	0.1	0.5	8	0.0	0.3
	B	694	3.2	17.7	550	2.6	14.7	67	0.3	2.1	8	0.0	0.4
	C	1370	6.4	34.9	1100	5.1	30.2	174	0.8	5.8	19	0.1	0.7
	D	1706	7.9	43.5	1320	6.1	35.6	652	3.0	21.9	97	0.5	4.1
Osteo-articular	A	333	1.5	21.5	297	1.4	19.4	88	0.4	6.2	16	0.1	1.2
	B	506	2.4	32.7	398	1.9	26.9	101	0.5	7.7	21	0.1	1.7
	C	349	1.6	22.5	299	1.4	19.8	112	0.5	8.7	28	0.1	2.4
	D	360	1.7	23.3	312	1.5	20.7	189	0.9	14.7	39	0.2	3.5
Endo-crine	A	78	0.4	12.1	62	0.3	9.9	15	0.1	2.8	3	0.0	0.6
	B	221	1.0	34.4	178	0.8	28.8	35	0.2	6.3	6	0.0	1.4
	C	160	0.7	24.9	134	0.6	21.4	41	0.2	7.7	7	0.0	1.8
	D	184	0.9	28.6	137	0.6	22.1	66	0.3	12.0	15	0.1	3.8
Heart	A	438	2.0	23.8	377	1.8	21.1	85	0.4	5.2	17	0.1	1.1
	B	647	3.0	35.2	478	2.2	27.1	126	0.6	8.3	14	0.1	1.0
	C	418	1.9	22.8	318	1.5	17.8	99	0.5	6.5	15	0.1	1.1
	D	334	1.6	18.2	231	1.1	13.1	134	0.6	9.0	42	0.2	3.3
Tumors	A	34	0.2	8.9	26	0.1	7.0	4	0.0	1.3	0	0.0	0.3
	B	116	0.5	30.3	55	0.3	16.0	3	0.0	1.5	0	0.0	0.0
	C	82	0.4	21.4	38	0.2	10.4	5	0.0	1.7	0	0.0	0.0
	D	151	0.7	39.4	35	0.2	9.5	6	0.0	1.9	2	0.0	0.7

(*continued on next page*)

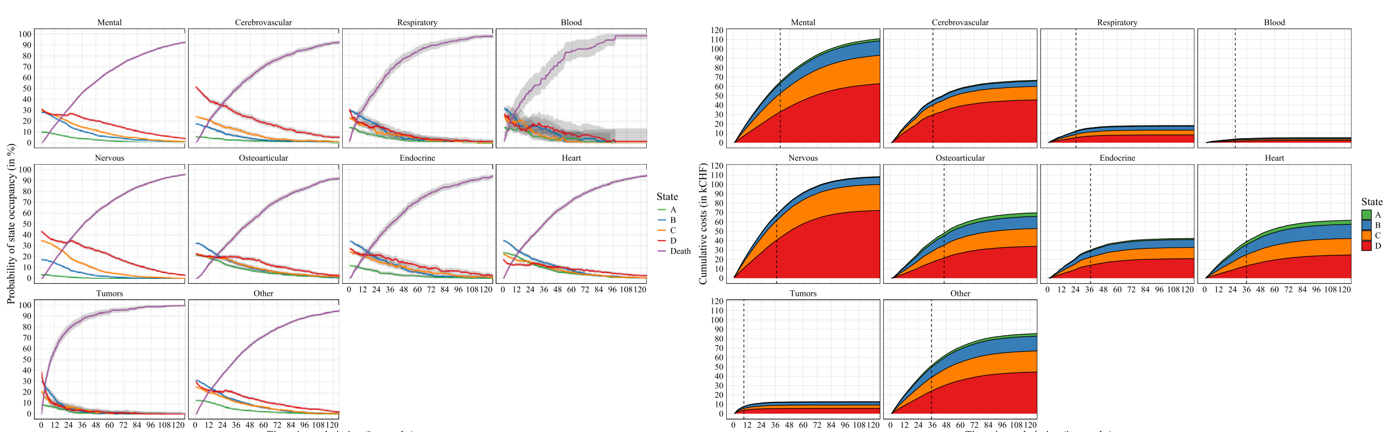
Table 5 (*continued*)

Variable	State	Entry M	%	$\hat{p}$	After 1 year M	%	$\hat{p}$	After 5 years M	%	$\hat{p}$	After 10 years M	%	$\hat{p}$
Other	A	532	2.5	13.0	466	2.2	11.7	105	0.5	2.9	23	0.1	0.7
	B	1284	6.0	31.4	987	4.6	25.1	213	1.0	6.5	39	0.2	1.4
	C	1038	4.8	25.4	784	3.6	20.3	220	1.0	6.9	33	0.2	1.2
	D	1235	5.7	30.2	870	4.0	22.2	374	1.7	12.0	73	0.3	3.1
Dependence in ADL DP													
1-6	A	1873	8.7	52.1	1570	7.3	44.9	356	1.7	11.6	70	0.3	2.6
	B	1661	7.7	46.2	1341	6.2	39.6	410	1.9	15.2	93	0.4	4.0
	C	56	0.3	1.6	165	0.8	4.9	245	1.1	9.3	63	0.3	2.9
	D	2	0.0	0.1	97	0.5	2.8	270	1.3	10.3	106	0.5	5.3
7	A	593	2.8	6.8	507	2.4	5.8	144	0.7	1.7	28	0.1	0.4
	B	4165	19.4	47.7	3146	14.6	37.3	488	2.3	6.5	74	0.3	1.1
	C	3319	15.4	38.0	2684	12.5	32.4	730	3.4	10.4	101	0.5	1.7
	D	661	3.1	7.6	877	4.1	10.5	1020	4.7	15.0	218	1.0	3.9
8	A	0	0.0	0.0	1	0.0	0.0	0	0.0	0.0	0	0.0	0.0
	B	97	0.5	1.3	93	0.4	1.3	25	0.1	0.4	1	0.0	0.0
	C	2651	12.3	35.1	1953	9.1	27.2	228	1.1	3.9	29	0.1	0.6
	D	4814	22.4	63.7	3398	15.8	47.2	1025	4.8	17.8	155	0.7	3.7
9	A	0	0.0	0.0	0	0.0	0.0	0	0.0	0.0	0	0.0	0.0
	B	0	0.0	0.0	1	0.0	0.1	0	0.0	0.1	0	0.0	0.1
	C	10	0.0	0.6	8	0.0	0.5	3	0.0	0.2	1	0.0	0.1
	D	1592	7.4	99.4	923	4.3	59.7	279	1.3	20.1	56	0.3	4.3
Physical mobility limitations PM													
1-5	A	1266	5.9	57.0	1107	5.2	50.6	309	1.4	15.3	62	0.3	3.4
	B	852	4.0	38.4	744	3.5	34.6	326	1.5	17.2	87	0.4	5.4
	C	87	0.4	3.9	133	0.6	6.2	175	0.8	9.4	55	0.3	3.6
	D	15	0.1	0.7	56	0.3	2.6	187	0.9	10.0	95	0.4	5.9
6	A	1110	5.2	17.7	894	4.2	14.8	171	0.8	3.2	32	0.1	0.7
	B	3119	14.5	49.8	2393	11.1	39.7	442	2.1	8.6	64	0.3	1.4
	C	1591	7.4	25.4	1395	6.5	23.6	495	2.3	10.0	69	0.3	1.7
	D	442	2.1	7.1	587	2.7	10.0	731	3.4	15.1	146	0.7	4.3
7	A	79	0.4	1.4	67	0.3	1.2	16	0.1	0.3	0	0.0	0.0
	B	1572	7.3	28.8	1190	5.5	22.9	122	0.6	3.0	16	0.1	0.5
	C	2586	12.0	47.4	2017	9.4	38.7	330	1.5	8.0	39	0.2	1.2
	D	1217	5.7	22.3	1066	5.0	20.5	660	3.1	16.5	104	0.5	3.4
8	A	11	0.1	0.5	9	0.0	0.4	4	0.0	0.2	3	0.0	0.2
	B	377	1.8	17.3	243	1.1	12.3	27	0.1	1.8	1	0.0	0.1
	C	1121	5.2	51.5	770	3.6	38.2	99	0.5	6.8	10	0.0	0.8
	D	667	3.1	30.7	519	2.4	25.3	196	0.9	13.4	28	0.1	2.5
9	A	0	0.0	0.0	1	0.0	0.0	0	0.0	0.0	1	0.0	0.0
	B	3	0.0	0.1	11	0.1	0.2	6	0.0	0.1	0	0.0	0.0
	C	651	3.0	12.1	495	2.3	9.6	107	0.5	2.4	21	0.1	0.5
	D	4728	22.0	87.8	3067	14.3	59.4	820	3.8	18.7	162	0.8	4.3
Orientation problems OR													
1-4	A	1574	7.3	45.1	1379	6.4	40.4	378	1.8	11.9	79	0.4	2.7
	B	1338	6.2	38.4	1080	5.0	32.1	407	1.9	13.5	98	0.5	3.6
	C	377	1.8	10.8	394	1.8	11.8	266	1.2	8.9	78	0.4	2.9
	D	199	0.9	5.7	225	1.0	6.7	306	1.4	10.4	119	0.6	4.6
5	A	876	4.1	11.7	684	3.2	9.4	117	0.5	1.9	18	0.1	0.3
	B	3417	15.9	45.8	2588	12.0	36.4	428	2.0	7.4	57	0.3	1.2
	C	1818	8.5	24.4	1483	6.9	20.8	509	2.4	9.1	74	0.3	1.8
	D	1345	6.3	18.0	1147	5.3	16.1	767	3.6	13.6	154	0.7	3.8
6	A	16	0.1	0.3	15	0.1	0.3	5	0.0	0.1	1	0.0	0.0
	B	1059	4.9	18.3	818	3.8	14.8	72	0.3	1.6	12	0.1	0.4
	C	2508	11.7	43.3	1894	8.8	34.9	280	1.3	6.6	21	0.1	0.6
	D	2212	10.3	38.2	1605	7.5	29.4	671	3.1	16.3	115	0.5	4.0
7	A	0	0.0	0.0	0	0.0	0.0	0	0.0	0.0	0	0.0	0.0
	B	103	0.5	4.3	88	0.4	3.7	16	0.1	0.7	1	0.0	0.0
	C	912	4.2	38.4	712	3.3	31.1	118	0.5	5.9	14	0.1	0.8
	D	1359	6.3	57.2	977	4.5	42.5	376	1.7	19.4	68	0.3	4.2
8-9	A	0	0.0	0.0	0	0.0	0.0	0	0.0	0.0	0	0.0	0.0
	B	6	0.0	0.3	7	0.0	0.3	0	0.0	0.0	0	0.0	0.0
	C	421	2.0	17.7	327	1.5	14.4	33	0.2	1.7	7	0.0	0.4
	D	1954	9.1	82.1	1341	6.2	58.6	474	2.2	24.6	79	0.4	4.7
Occupational limitations OC													
1-5	A	1139	5.3	56.6	1004	4.7	50.6	286	1.3	15.3	65	0.3	3.7
	B	743	3.5	36.9	648	3.0	33.0	272	1.3	15.2	79	0.4	4.8
	C	100	0.5	5.0	139	0.6	7.1	171	0.8	9.6	55	0.3	3.3
	D	31	0.1	1.5	60	0.3	3.0	191	0.9	10.8	84	0.4	5.4
6	A	1042	4.8	17.7	833	3.9	14.6	157	0.7	3.2	22	0.1	0.6
	B	2883	13.4	49.0	2225	10.4	39.6	408	1.9	8.9	57	0.3	1.5
	C	1334	6.2	22.7	1167	5.4	20.7	423	2.0	9.4	63	0.3	1.9
	D	630	2.9	10.7	683	3.2	12.1	602	2.8	13.4	127	0.6	4.0

Table 5 (*continued*)

Variable	State	Entry M	%	$\hat{p}$	After 1 year M	%	$\hat{p}$	After 5 years M	%	$\hat{p}$	After 10 years M	%	$\hat{p}$
7	A	255	1.2	2.5	215	1.0	2.2	47	0.2	0.5	10	0.0	0.1
	B	2137	9.9	21.1	1570	7.3	16.3	207	1.0	2.7	29	0.1	0.5
	C	3853	17.9	38.1	2912	13.5	30.7	481	2.2	6.5	52	0.2	0.9
	D	3875	18.0	38.3	2854	13.3	30.0	1158	5.4	16.0	190	0.9	3.7
8-9	A	30	0.1	0.9	26	0.1	0.7	10	0.0	0.3	1	0.0	0.0
	B	160	0.7	4.6	138	0.6	4.0	36	0.2	1.1	3	0.0	0.1
	C	749	3.5	21.6	592	2.8	17.5	131	0.6	4.1	24	0.1	0.8
	D	2533	11.8	73.0	1698	7.9	50.2	643	3.0	20.9	134	0.6	4.7
Social integration limitations SI													
1-4	A	921	4.3	48.8	818	3.8	44.2	220	1.0	12.8	45	0.2	2.9
	B	667	3.1	35.4	566	2.6	30.7	215	1.0	13.1	55	0.3	3.7
	C	197	0.9	10.4	202	0.9	11.0	164	0.8	10.1	42	0.2	2.9
	D	101	0.5	5.4	109	0.5	6.0	158	0.7	9.9	69	0.3	5.1
5	A	1172	5.5	22.7	948	4.4	18.9	202	0.9	4.5	42	0.2	1.0
	B	2379	11.1	46.1	1842	8.6	37.3	412	1.9	9.8	70	0.3	2.0
	C	964	4.5	18.7	887	4.1	18.1	353	1.6	8.6	70	0.3	2.2
	D	642	3.0	12.4	620	2.9	12.4	539	2.5	13.0	121	0.6	3.8
6	A	361	1.7	4.8	301	1.4	4.0	75	0.3	1.2	10	0.0	0.2
	B	2423	11.3	31.9	1828	8.5	25.3	247	1.1	4.3	35	0.2	0.7
	C	2675	12.4	35.2	2066	9.6	28.6	438	2.0	7.8	52	0.2	1.1
	D	2140	10.0	28.2	1617	7.5	22.4	827	3.8	14.8	155	0.7	3.8
7	A	12	0.1	0.2	11	0.1	0.2	3	0.0	0.1	1	0.0	0.0
	B	452	2.1	8.5	343	1.6	6.7	49	0.2	1.2	8	0.0	0.3
	C	2025	9.4	38.0	1517	7.1	30.4	232	1.1	5.8	24	0.1	0.7
	D	2838	13.2	53.3	2028	9.4	40.3	718	3.3	18.4	117	0.5	4.0
8-9	A	0	0.0	0.0	0	0.0	0.0	0	0.0	0.0	0	0.0	0.0
	B	2	0.0	0.1	2	0.0	0.1	0	0.0	0.1	0	0.0	0.1
	C	175	0.8	11.5	138	0.6	9.2	19	0.1	1.3	6	0.0	0.4
	D	1348	6.3	88.4	921	4.3	62.0	352	1.6	26.2	73	0.3	5.9

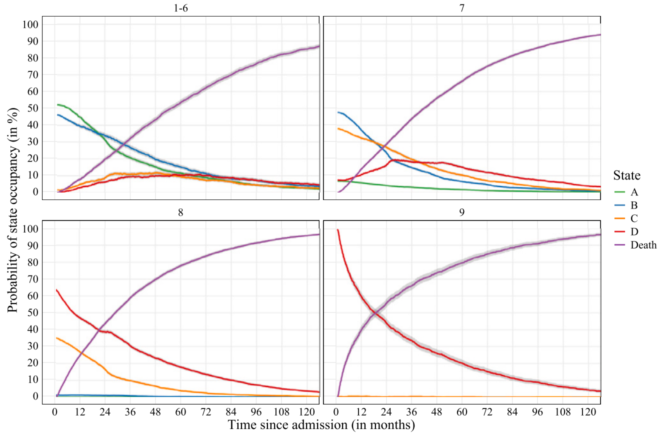
Note: Column "M" represents the number of individuals at risk, and "%" indicates the share calculated on the original 21 494 individuals. The Aalen-Johansen estimate of the occupation probability is presented in "$\hat{p}$".



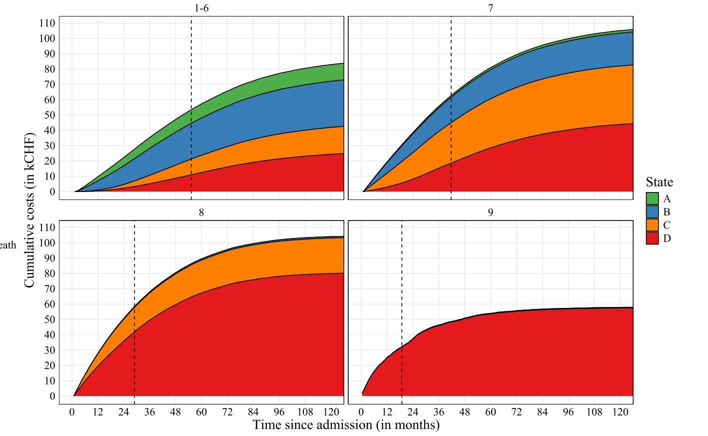
(a) Aalen-Johansen estimates. (b) Cumulative costs.

Note: See Fig. 5.

Fig. 11. Aalen-Johansen estimates with 95% confidence intervals of state occupancy probabilities and cumulative 10-year LTC costs stratified by the primary medical diagnosis ($D1$).



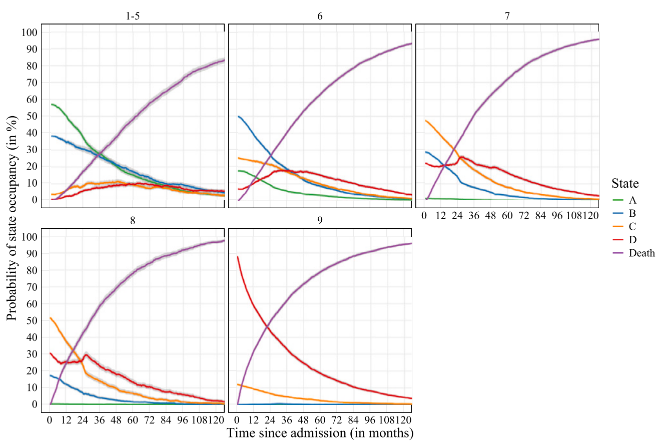
(a) Aalen-Johansen estimates.



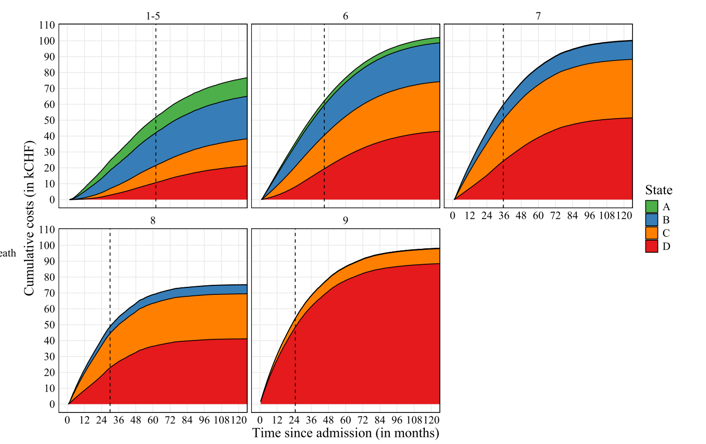
(b) Cumulative costs.

Note: See Fig. 5.

Fig. 12. Aalen-Johansen estimates with 95% confidence interval of state occupancy probabilities and cumulative 10-year LTC costs stratified by the dependency from others (DP).



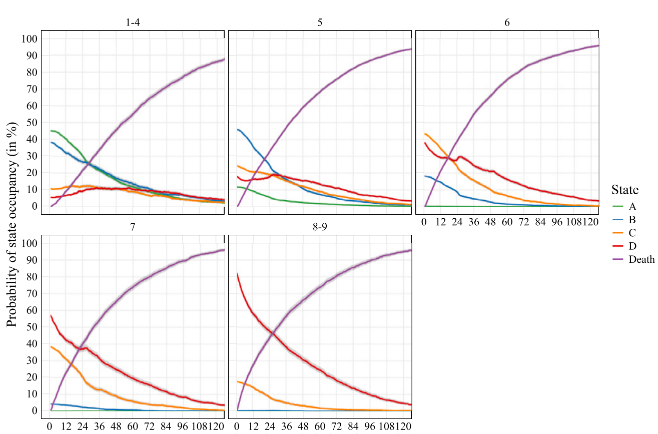
(a) Aalen-Johansen estimates.



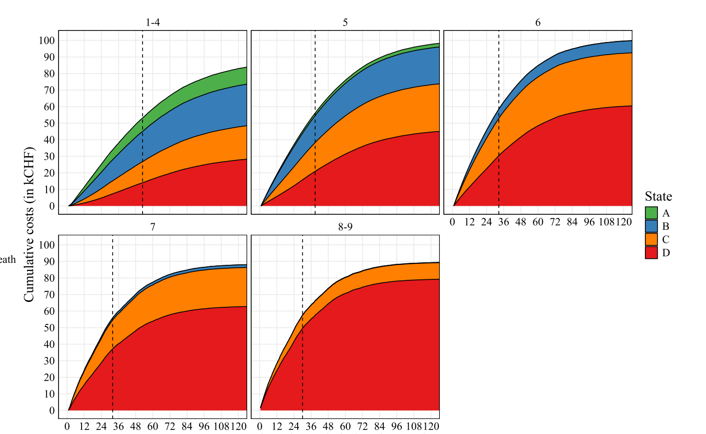
(b) Cumulative costs.

Note: See Fig. 5.

Fig. 13. Aalen-Johansen estimates with 95% confidence interval of state occupancy probabilities and cumulative 10-year LTC costs stratified by the physical mobility (PM).



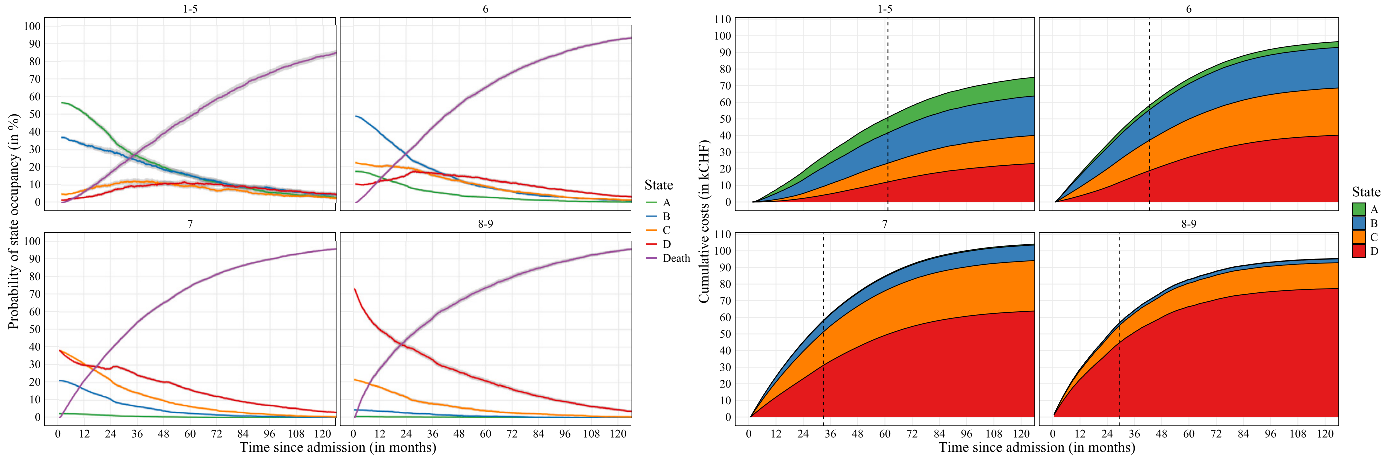
(a) Aalen-Johansen estimates.



(b) Cumulative costs.

Note: See Fig. 5.

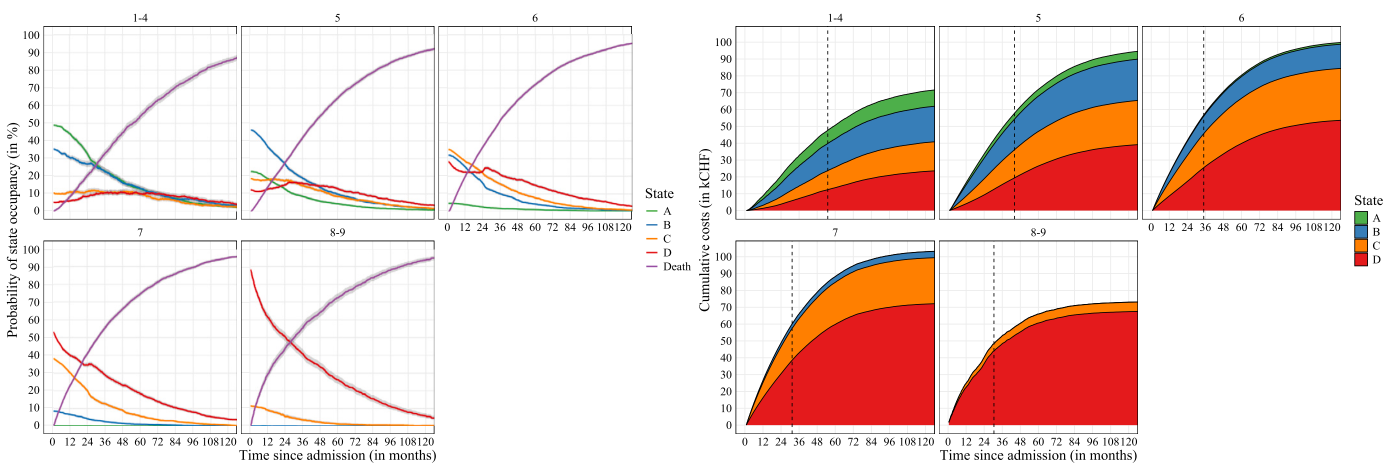
Fig. 14. Aalen-Johansen estimates with 95% confidence interval of state occupancy probabilities and cumulative 10-year LTC costs stratified by the orientation in space (OR).



(a) Aalen-Johansen estimates. (b) Cumulative costs.

Note: See Fig. 5.

Fig. 15. Aalen-Johansen estimates with 95% confidence interval of state occupancy probabilities and cumulative 10-year LTC costs stratified by the occupation (OC).



(a) Aalen-Johansen estimates. (b) Cumulative costs.

Note: See Fig. 5.

Fig. 16. Aalen-Johansen estimates with 95% confidence interval of state occupancy probabilities and cumulative 10-year LTC costs stratified by the social integration (SI).

Data availability

The authors do not have permission to share data.

References

Ai, J., Brockett, P.L., Golden, L.L., Zhu, W., 2017. Health state transitions and longevity effects on retirees' optimal annuitization. J. Risk Insur. 84 (S1), 319–343.

Akaike, H., 1974. A new look at the statistical model identification. IEEE Trans. Autom. Control 19 (6), 716–723.

Albarrán, I., Alonso-González, P.J., Arribas-Gil, A., Grané, A., 2019. How functional data can enhance the estimation of health expectancy: the case of disabled Spanish population. ASTIN Bull. 49 (1), 57–84.

Arrighi, H.M., McLaughlin, T., Leibman, C., 2010. Prevalence and impact of dementia-related functional limitations in the United States, 2001 to 2005. Alzheimer Dis. Assoc. Disord. 24 (1).

Barnett, K., Mercer, S.W., Norbury, M., Watt, G., Wyke, S., Guthrie, B., 2012. Epidemiology of multimorbidity and implications for health care, research, and medical education: a cross-sectional study. Lancet 380 (9836), 37–43.

Biessy, G., 2015. Long-term care insurance: a multi-state semi-markov model to describe the dependency process for elderly people. Bull. Fr. Actuariat 15 (29), 41–74.

Bladt, M., Fuino, M., Shemendyuk, A., Wagner, J., 2023. Modelling the burden of long-term care for institutionalised elderly based on care duration and intensity. Ann. Actuar. Sci. 17 (1), 83–117.

Börsch-Supan, A., Brandt, M., Hunkler, C., Kneip, T., Korbmacher, J., Malter, F., Schaan, B., Stuck, S., Zuber, S., 2013. Data resource profile: the survey of health, ageing and retirement in Europe (SHARE). Int. J. Epidemiol. 42 (4), 992–1001.

Brandis, M., Stacom, R., 2009. Long-term care in the home for people with multiple sclerosis. Care Manag. J. 10 (3), 128–137.

Brown, J.R., Finkelstein, A., 2009. The private market for long-term care insurance in the United States: a review of the evidence. J. Risk Insur. 76 (1), 5–29.

Burke, O., Levine, S.K., Mir, N., Gleason, L.J., 2023. Transitions of care for advanced cancer in post-acute and long-term care. J. Am. Geriatr. Soc. 71 (5), 1680–1684.

Burt, J., Rick, J., Blakeman, T., Protheroe, J., Roland, M., Bower, P., 2014. Care plans and care planning in long-term conditions: a conceptual model. Primary Health Care Res. Develop. 15 (4), 342–354.

Chan, A., Zimmer, Z., Saito, Y., 2011. Gender differentials in disability and mortality transitions: the case of older adults in Japan. J. Aging Health 23 (8), 1285–1308. PMID: 21653794.

Christiansen, M.C., 2012. Multistate models in health insurance. AStA Adv. Stat. Anal. 96 (2), 155–186.

Colombo, F., Llena-Nozal, A., Mercier, J., Tjadens, F., 2011. Help Wanted? Providing and Paying for Long-Term Care. OECD Health Policy Studies. OECD Publishing, Paris.

Cosandey, J., 2016. De Nouvelles Mesures pour les Soins aux Personnes Âgées. Avenir Suisse, Zurich, Switzerland.

Cox, D.R., Miller, H.D., 1965. The Theory of Stochastic Processes. Chapman and Hall, London.

Czado, C., Rudolph, F., 2002. Application of survival analysis methods to long-term care insurance. Insur. Math. Econ. 31 (3), 395–413.

Da Roit, B., Le Bihan, B., 2010. Similar and yet so different: cash-for-care in six European countries' long-term care policies. Milbank Q. 88 (3), 286–309.

Davidson, I.A., Dewey, M.E., Copeland, J.R.M., 1988. The relationship between mortality and mental disorder: evidence from the liverpool longitudinal study. Int. J. Geriatr. Psychiatry 3 (2), 95–98.
Deeg, D.J.H., Portrait, F., Lindeboom, M., 2002. Health profiles and profile-specific health expectancies of older women and men: the Netherlands. J. Women Aging 14 (1–2), 27–46.
Deshmukh, S., 2012. Multi-State Transition Models for Cash Flows. Springer India, India, pp. 173–203.
Dorr, D.A., Horn, S.D., Smout, R.J., 2005. Cost analysis of nursing home registered nurse staffing times. J. Am. Geriatr. Soc. 53 (5), 840–845.
Esquível, M.L., Guerreiro, G.R., Oliveira, M.C., Corte Real, P., 2021. Calibration of transition intensities for a multistate model: application to long-term care. Risks 9 (2), 37.
Federal Department of Home Affairs, 2016. Ordonnance du DFI du 29 septembre 1995 sur les prestations dans l'assurance obligatoire des soins en cas de maladie. In French. Accessed: 20 February, 2024.
Fisher, G.G., Ryan, L.H., 2017. Overview of the health and retirement study and introduction to the special issue. Work Aging Retir. 4 (1), 1–9.
Fong, J.H., 2019. Disability incidence and functional decline among older adults with major chronic diseases. BMC Geriatr. 19 (1), 323.
Fong, J.H., Shao, A.W., Sherris, M., 2015. Multistate actuarial models of functional disability. N. Am. Actuar. J. 19 (1), 41–59.
Fong, J.H., Sherris, M., Yap, J., 2017. Forecasting disability: application of a frailty model. Scand. Actuar. J. 2017 (2), 125–147.
Foucher, Y., Giral, M., Soulillou, J., Daures, J., 2010. A flexible semi-Markov model for interval-censored data and goodness-of-fit testing. Stat. Methods Med. Res. 19 (2), 127–145. Publisher: SAGE Publications Ltd STM.
Freedman, V.A., Wolf, D.A., Spillman, B.C., 2016. Disability-free life expectancy over 30 years: a growing female disadvantage in the us population. Am. J. Publ. Health 106 (6), 1079–1085. PMID: 26985619.
Fuino, M., Wagner, J., 2018. Long-term care models and dependence probability tables by acuity level: new empirical evidence from Switzerland. Insur. Math. Econ. 81, 51–70.
Fuino, M., Wagner, J., 2020. Duration of long-term care: socio-economic factors, type of care interactions, and evolution. Insur. Math. Econ. 90, 151–168.
Germain, C.M., Vasquez, E., Batsis, J.A., McQuoid, D.R., 2016. Sex, race, and age differences in muscle strength and limitations in community dwelling older adults: data from the health and retirement survey (HRS). Arch. Gerontol. Geriatr. 65, 98–103.
Govorun, M., Latouche, G., Loisel, S., 2015. Phase-type aging modeling for health dependent costs. Insur. Math. Econ. 62, 173–183.
Guccione, A.A., Felson, D.T., Anderson, J.J., Anthony, J.M., Zhang, Y., Wilson, P.W., Kelly-Hayes, M., Wolf, P.A., Kreger, B.E., Kannel, W.B., 1994. The effects of specific medical conditions on the functional limitations of elders in the framingham study. Am. J. Publ. Health 84 (3), 351–358.
Haberman, S., Pitacco, E., 2018. Actuarial Models for Disability Insurance. Routledge, New York.
Haller, G., Combescure, C., Mamie, C., Zoccatelli, D., Clergue, F., 2015. Impact of policies regulating foreign physician migration to Switzerland: a modelling case study in anaesthesia. BMC Health Serv. Res. 15 (1), 204.
Helms, F., Czado, C., Gschlößl, S., 2005. Calculation of LTC premiums based on direct estimates of transition probabilities. ASTIN Bull. 35 (2), 455–469.
Hill, M., Lambert, P.C., Crowther, M.J., 2021. Relaxing the assumption of constant transition rates in a multi-state model in hospital epidemiology. BMC Med. Res. Methodol. 21 (1), 16.
Jackson, C., 2011. Multi-state models for panel data: the msm package for R. J. Stat. Softw. 38, 1–28.
Jacobs, B.J., 2012. Issues in Caregiving for Cancer Patients with Long-Term Care Needs. Springer New York, New York, NY, pp. 145–159.
Janssen, J., Manca, R., 2001. Numerical solution of non-homogeneous semi-Markov processes in transient case*. Methodol. Comput. Appl. Probab. 3 (3), 271–293.
Jennings, L.A., Hollands, S., Keeler, E., Wenger, N.S., Reuben, D.B., 2020. The effects of dementia care co-management on acute care, hospice, and long-term care utilization. J. Am. Geriatr. Soc. 68 (11), 2500–2507.
Juster, F.T., Suzman, R., 1995. An overview of the health and retirement study. J. Hum. Resour. 30, S7–S56.
Kalaria, R.N., 2012. Cerebrovascular disease and mechanisms of cognitive impairment. Stroke 43 (9), 2526–2534.
Kalbfleisch, J.D., Lawless, J.F., 1985. The analysis of panel data under a markov assumption. J. Am. Stat. Assoc. 80 (392), 863–871.
Katz, P.R., 2011. An international perspective on long term care: focus on nursing homes. J. Am. Med. Dir. Assoc. 12 (7). 487–492.e1.
Kay, R., 1986. A Markov model for analysing cancer markers and disease states in survival studies. Biometrics 42 (4), 855–865.
Kempen, G., Myers, A.M., Powell, L.E., 1995. Hierarchical structure in ADL and IADL: analytical assumptions and applications for clinicians and researchers. J. Clin. Epidemiol. 48 (11), 1299–1305.
Kingston, A., Davies, K., Collerton, J., Robinson, L., Duncan, R., Bond, J., Kirkwood, T.B.L., Jagger, C., 2014. The contribution of diseases to the male-female disability-survival paradox in the very old: results from the newcastle 85+ study. PLoS ONE 9 (2), 1–10.
Koroukian, S.M., Schiltz, N., Warner, D.F., Sun, J., Bakaki, P.M., Smyth, K.A., Stange, K.C., Given, C.W., 2016. Combinations of chronic conditions, functional limitations, and geriatric syndromes that predict health outcomes. J. Gen. Intern. Med. 31 (6), 630–637.
Liddle, P.F., 1992. Severe persistent mental illness. Psych. Bull. 16 (12), 743–745.
Lin, C.-C., 2017. When cancer care becomes a long-term care issue: are we ready? Cancer Nurs. 40 (5).
Makam, A.N., Tran, T., Miller, M.E., Xuan, L., Nguyen, O.K., Halm, E.A., 2019. The clinical course after long-term acute care hospital admission among older medicare beneficiaries. J. Am. Geriatr. Soc. 67 (11), 2282–2288.
Marengoni, A., Angleman, S., Melis, R., Mangialasche, F., Karp, A., Garmen, A., Meinow, B., Fratiglioni, L., 2011. Aging with multimorbidity: a systematic review of the literature. Ageing Res. Rev. 10 (4), 430–439.
Marshall, G., Jones, R.H., 1995. Multi-state models and diabetic retinopathy. Stat. Med. 14 (18), 1975–1983.
Mathers, C., 1996. Trends in health expectancies in Australia 1981–1993. J. Austral. Popul. Assoc. 13 (1), 1–15.
Nichols, B.L., Davis, C.R., Richardson, D.R., 2010. An integrative review of global nursing workforce issues. Annu. Rev. Nurs. Res. 28, 113–132.
Nihtilä, E.K., Martikainen, P.T., Koskinen, S.V., Reunanen, A.R., Noro, A.M., Häkkinen, U.T., 2007. Chronic conditions and the risk of long-term institutionalization among older people. Eur. J. Public Health 18 (1), 77–84.
OECD, 2017. Health at a Glance 2017: OECD Indicators. OECD Publishing, Paris.
Pack, S., 2009. Poor physical health and mortality in patients with Schizophrenia. Nurs. Stand. (Royal College of Nursing Great Britain, 1987) 23 (21), 41–45.
Pinheiro, J.C., Bates, D.M., 2000. Mixed-effects models in S and S-PLUS. In: Statistics and Computing, 1 edition. Springer New York, NY, New York, NY.
Pritchard, D.J., 2006. Modeling disability in long-term care insurance. N. Am. Actuar. J. 10 (4), 48–75.
Rabiner, L., 1989. A tutorial on hidden markov models and selected applications in speech recognition. Proc. IEEE 77 (2), 257–286.
Republic and Canton of Geneva, General Directorate for Health, 2019. EROS/PLAISIR data from 1996-2018.
Rickayzen, B.D., Walsh, D.E.P., 2002. A multi-state model of disability for the United Kingdom: implications for future need for long-term care for the elderly. Br. Actuar. J. 8 (2), 341–393.
Roussel, B., Tilquin, C., 1993. Bio-Psycho-Social Profile of Clients in Long Term Care and Planning of Nursing Care Required. Équipe de Recherche Opérationnelle en Santé, Montreal. Reference manual.
Shemendyuk, A., Wagner, J., 2024. On the factors determining the health profiles and care needs of institutionalized elders. Insur. Math. Econ. 114, 223–241.
Sherris, M., Wei, P., 2021. A multi-state model of functional disability and health status in the presence of systematic trend and uncertainty. N. Am. Actuar. J. 25 (1), 17–39.
Sonnega, A., Faul, J.D., Ofstedal, M.B., Langa, K.M., Phillips, J.W.R., Weir, D.R., 2014. Cohort profile: the health and retirement study (HRS). Int. J. Epidemiol. 43 (2), 576–585.
Stark, A.J., Kane, R.L., Kane, R.A., Finch, M., 1995. Effect on physical functioning of care in adult Foster homes and nursing homes. Gerontologist 35 (5), 648–655.
Tamiya, N., Kobayashi, Y., Murakami, S., Sasaki, J., Yoshizawa, K., Otaki, J., Kano, K., 2001. Factors related to home discharge of cerebrovascular disease patients: 1-year follow-up interview survey of caregivers of hospitalized patients in 53 acute care hospitals in Japan. Arch. Gerontol. Geriatr. 33 (2), 109–121.
Therneau, T.M., 2024. A Package for Survival Analysis in R. R package version 3.5-8.
Titman, A.C., 2011. Flexible nonhomogeneous Markov models for panel observed data. Biometrics 67 (3), 780–787.
Turner-Stokes, L., Sykes, N., Silber, E., 2008. Long-term neurological conditions: management at the interface between neurology, rehabilitation and palliative care. Clin. Med. 8 (2), 186–191.
van den Akker, M., Buntinx, F., Metsemakers, J.F.M., Roos, S., Knottnerus, J., 1998. Multimorbidity in general practice: prevalence, incidence, and determinants of co-occurring chronic and recurrent diseases. J. Clin. Epidemiol. 51 (5), 367–375.
World Health Organization, 1980. International classification of impairments, disabilities, and handicaps: a manual of classification relating to the consequences of disease, published in accordance with resolution WHA29.35 of the Twenty-ninth World Health Assembly, May 1976.
World Health Organization, 2016. International Statistical Classification of Diseases and Related Health Problems - 10th revision. Geneva, Switzerland, Fifth edition.
Xu, X., Chan, Y.H., Chan, Q.L., Gyanwali, B., Hilal, S., Tan, B.Y., Ikram, M.K., Venketasubramanian, N., Chen, C.L.-H., 2018. Global cerebrovascular burden and long-term clinical outcomes in asian elderly across the spectrum of cognitive impairment. Int. Psychogeriatr. 30 (9), 1355–1363.
Yang, W., Jingwei He, A., Fang, L., Mossialos, E., 2016. Financing institutional long-term care for the elderly in China: a policy evaluation of new models. Health Policy Planning 31 (10), 1391–1401.
Zhao, Y., Hu, Y., Smith, J.P., Strauss, J., Yang, G., 2012. Cohort profile: the China health and retirement longitudinal study (CHARLS). Int. J. Epidemiol. 43 (1), 61–68.
Zúñiga, F., Jenni, G., Wiesli, U., Schwendimann, R., 2010. Entwicklung der Rolle der Advanced Practice Nurse in der stationären Langzeitpflege älterer Menschen in der Schweiz. Pflege 23 (6), 375–383. Publisher: Hogrefe AG.